

**An exploration of therapist and dose effects within an exercise  
intervention study for reducing falls and falls risk factors in community  
dwelling older people.**

Sheena Gawler

UCL

Doctor of Philosophy

**Declaration**

I, Sheena Gawler, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

## **Abstract**

### **Background:**

Therapist and dose effects may be important in falls prevention exercise service delivery because the evidence for falls prevention exercise has a very specific prescription, therefore, maximising therapist skills, minimising 'therapist drift' and encouraging compliance could enhance patient outcomes.

### **Methods:**

The aims of this research were to study (a) the effect of the therapist (delivering the exercise programme), and (b) the effect of exercise dose, on falls outcomes within a group exercise intervention. The primary objective was to establish any difference in the number of falls for subjects participating in the intervention and the secondary objective was to establish any difference in falls risk factors (balance and lower limb power) for these subjects, according to their therapist and separately, their dose. Multilevel modelling, which is designed for clustered data, was used to investigate the magnitude of therapist and dose effects, and to explore whether specific therapist characteristics were individually associated with the falls outcomes.

### **Results and Conclusion:**

Unconditional multilevel models showed some variance between patients grouped by therapists of up to 6% of the overall variance in falls outcomes. These effect sizes are small, but in a standardised exercise intervention, especially within the research setting, they would not be expected to be large due to quality assurance procedures reducing variability between therapists. The therapist characteristics investigated, however, did not explain this therapist-level variance, and it may be that the characteristics studied did not include those that make a difference to falls outcomes. Another explanation for the unconvincing evidence of therapists effects is that the main trial (within which my study was nested) was not set up to investigate therapist effects and therefore was not powered for this. The dose effect analysis showed that the dose of the exercise intervention was not an independent predictor of falls rate nor falls risk factors. It is possible that the dose investigation was affected by the high numbers of non-fallers within the recruited population. The use of our protocol and

documents for the quality assurance of the intervention within research was effective at standardisation and ensuring fidelity, and this approach could be used as part of falls prevention exercise service delivery to reduce 'therapist drift'.

## Table of Contents

TITLE .....	1
DECLARATION.....	2
ABSTRACT .....	3
LIST OF TABLES .....	10
LIST OF FIGURES .....	12
LIST OF APPENDICES .....	13
LIST OF ABBREVIATIONS AND SYMBOLS.....	14
ACKNOWLEDGEMENTS .....	16
CHAPTER 1.....	17
Background.....	17
1.1 Chapter summary .....	17
1.2 Definition of a fall.....	17
1.3 Why study falls? .....	20
1.3.1 Prevalence of falls .....	20
1.3.2 Consequences of falls.....	21
1.3.2.1 Financial cost of falls.....	21
1.3.2.2 Consequences of falls for individuals .....	22
1.4 Risk factors for falling .....	23
1.5 Identifying those at risk .....	27
1.6 Monitoring and reporting falls .....	30
1.7 Outcome measures.....	33
1.8 Falls prevention exercise research .....	35
1.8.1 Falls prevention exercise prescription.....	36
1.8.2 Walking .....	36
1.9 Clinical practice documents .....	37
1.10 Falls exercise provision .....	38
1.10.1 Falls interventions: Otago Home Exercise Programme .....	39
1.10.2 Falls interventions: Falls Management Exercise .....	45
1.11 Why study therapist effects?.....	46
1.11.1 Definition and purpose.....	46
1.12 Why study dose? .....	49
1.13 Research questions and rationale .....	51
1.13.1 Aims, objectives and hypotheses .....	52
1.14 Chapter conclusions.....	56
CHAPTER 2.....	57
Literature Review .....	57
2.1 Chapter summary .....	57
2.2 Rationale for literature review.....	57
2.3 Methods.....	59
2.3.1 Search Strategy .....	59
2.3.2 Data Extraction .....	61

2.3.3 Data analysis and synthesis .....	62
2.4 Results .....	62
2.4.1 Study selection.....	62
2.4.2 Study Patient Population.....	63
2.4.3 Intervention/therapy .....	71
2.4.4 Therapists .....	71
2.4.5 Therapist effect analysed .....	77
2.4.8 Results of the narrative review.....	82
2.4.8.1 Therapists' years of experience.....	85
2.4.8.2 Therapists' advanced/specialist qualification .....	88
2.4.8.3 Other therapist effects .....	90
2.4.8.4 Therapist effects in group exercise.....	91
2.5 Discussion of findings .....	93
2.5.1 Summary of findings .....	93
2.5.2 Findings in context of other therapist effects literature .....	93
2.5.3 Meaning and implications .....	94
2.5.4 Strengths and limitations of the review .....	97
2.6 Chapter conclusions.....	98
CHAPTER 3.....	99
Methods .....	99
Statement of Intellectual Property (IP) .....	99
3.1 Chapter summary .....	99
3.2 ProAct65+.....	100
3.2.1 Design, aims and objectives.....	100
3.2.2 FaME and OEP interventions .....	101
3.2.2.1 ProAct65+ Otago Home Exercise Programme .....	101
3.2.2.2 ProAct65+ Falls Management Exercise (FaME) programme .....	103
3.2.2.3 Usual Care .....	104
3.2.3 Sample size .....	104
3.2.4 Recruitment of general practices.....	104
3.2.5 Patient selection, recruitment and consenting.....	105
3.2.5.1 Recruitment of participants .....	105
3.2.5.2 Recruitment procedure.....	106
3.2.6 Allocation concealment and randomisation.....	106
3.2.7 Blinding .....	107
3.2.9 Recruitment and training of peer mentors and PSIs .....	111
3.2.9.1 Recruitment and training of PSIs.....	111
3.2.9.2 Recruitment and training of PMs .....	112
3.2.10 Quality control in the intervention arms.....	113
3.2.10.1 FaME quality control summary .....	113
3.2.10.2 OEP quality control.....	114
3.2.10.3 Patient intervention evaluations.....	114
3.3 Falls.....	115
3.3.1 Aims & objectives .....	115
3.3.2 Falls outcomes & falls data collection .....	115
3.3.2.1 Falls numbers & falls diaries .....	116

3.3.2.2 Functional assessments & falls risk assessment.....	119
3.3.2.2.1 Timed-Up-and-Go .....	119
3.3.2.2.2 Falls Risk Assessment Tool .....	120
3.3.2.2.3 Functional Reach.....	121
3.3.2.2.4 Modified Clinical Romberg Static Balance test .....	121
3.3.2.2.5 30 Second Chair Rise .....	122
3.3.2.3 Fear of falling.....	122
3.3.3 Data management .....	123
3.3.4 Baseline falls and falls risk analysis.....	124
3.3.5 Falls analysis .....	125
3.4 The therapist and dose effects study.....	126
3.4.1 Aims .....	126
3.4.2 Patient selection .....	126
3.4.3 Sample size .....	126
3.4.4 Falls & falls risk factors outcome measures.....	127
3.4.5 Data collection and management .....	127
3.4.5.1 Postural Stability Instructor (PSI) variables.....	127
3.4.5.2 FaME Quality Assurance (QA) visits.....	127
3.4.5.3 Use of quality control data to calculate therapist ‘quality’ categorisation.....	132
3.4.5.4 Patient evaluations.....	133
3.4.5.5 Patient attendance (dose) .....	134
3.4.5.6 Data entry .....	134
3.4.6 Blinding .....	134
3.4.7 Therapist and dose effects analysis plan 3.4.7.1 Data cleaning .....	135
3.4.7.2 Baseline characterisation.....	135
3.4.7.3 Bivariate analyses.....	135
3.4.7.4 Multiple regression equations .....	137
3.4.7.5 Multilevel modelling adjusting for baseline clustering.....	138
3.4.7.6 Multiple testing .....	139
3.5 Chapter conclusions.....	139
CHAPTER 4.....	140
Results 1; ProAct65+ baseline characteristics & ProAct65+ falls analyses .....	140
4.1 Chapter summary .....	140
4.2 ProAct65+ recruitment .....	140
4.3 Baseline ProAct65+ patient demographic characteristics.....	140
4.4 Baseline ProAct65+ patient falls characteristics.....	141
4.4.1 Protocol violations .....	141
4.4.2 Baseline falls characteristics .....	141
4.4.3 ProAct65+ population falls compared with UK older population .....	142
4.4.4 Baseline falls characteristics by study group .....	144
4.5 Functional assessment outcomes at baseline .....	145
4.6 ProAct65+ continuation of exercise.....	146
4.7 Adherence .....	147
4.8 Falls; Intervention and Follow Up.....	147
4.9 Falls diary data .....	151

4.10 Falls risk factors .....	152
4.11 Chapter conclusions.....	154
<b>CHAPTER 5.....</b>	<b>155</b>
Results 2; Dose and therapist effects.....	155
5.1 Chapter summary .....	155
5.2 Therapist characterisation .....	155
5.3 Baseline characteristics of therapist effect study population .....	156
5.4 Participant evaluation .....	159
5.5 Sample size.....	161
5.6 Therapist and dose effects analysis results.....	162
5.6.1 Data cleaning .....	162
5.6.2 Bivariate analyses .....	162
5.6.2.1 Correlation analyses .....	170
5.6.2.1.1 Exercise dose .....	170
5.6.2.1.2 Therapist Characteristics; PSI age & PSI attendance.....	176
5.6.2.3 Analysis of variance (ANOVA); therapist characteristics .....	190
5.6.2.3.1 Patient enjoyment of classes.....	192
5.6.2.3.2. Patient-rated intensity of classes.....	193
5.6.2.4 Summary; bivariate analyses .....	193
5.6.3 Multiple regression equations.....	195
5.6.4 Multilevel modelling .....	195
5.6.4.1 Unconditional models.....	195
5.6.4.2 Adding predictors to the models.....	196
5.6.4.2.1 Timed Up & Go (TUG).....	197
5.6.4.2.2 Chair rise .....	203
5.6.4.2.3 Romberg.....	205
5.6.4.2.4 Falls rate during the intervention .....	207
5.6.4.3 Summary; multilevel modelling .....	208
5.7 Chapter conclusions.....	211
<b>CHAPTER 6.....</b>	<b>212</b>
Discussion.....	212
6.1 Chapter summary .....	212
6.2 Summary of findings.....	212
6.2.1 Falls prevention .....	212
6.2.2 Therapist effect study; bivariate analyses.....	213
6.2.3 Therapist effect study; multilevel modelling analyses .....	214
6.2.3.1 Multilevel modelling; unconditional models .....	214
6.2.3.2 Multilevel modelling; adding predictors to the models.....	215
6.3 Relationship to other research.....	216
6.3.1 Falls prevention .....	216
6.3.2 Therapist effects .....	218
6.3.2.1 Standardised interventions.....	218
6.3.2.2 Therapist characteristics studied.....	219
6.3.2.3 Functional assessments .....	221
6.3.2.4 Quality assurance of intervention .....	222
6.3.2.5 Clinical experience.....	223



6.3.2.6 Time-points studied.....	224
6.3.2.7 Gender .....	225
6.3.3 Adherence & dose.....	225
6.4 Limitations and strengths .....	226
6.5 Implications for service provision.....	230
6.5.1 Clinical practice versus research .....	230
6.5.2 Quality assurance in clinical practice .....	233
6.6 Implications for further research.....	235
6.7 Implications for training of exercise professionals .....	236
6.8 Conclusions .....	237
REFERENCES .....	238
APPENDICES .....	255

## List of tables

<b>Table 2.1</b> , Study description; RCTs.....	66
<b>Table 2.2</b> , Study description; qualitative studies.....	67
<b>Table 2.3</b> , Study description; cohort studies .....	68
<b>Table 2.4</b> , Study description; case-control studies.....	70
<b>Table 2.5</b> , Therapy type .....	73
<b>Table 2.6</b> , Therapist characteristics .....	75
<b>Table 2.7</b> , Methodological quality of RCTs .....	79
<b>Table 2.8</b> , Methodological quality of qualitative studies .....	79
<b>Table 2.9</b> , Methodological quality of cohort studies.....	80
<b>Table 2.10</b> , Methodological quality of case-control studies .....	80
<b>Table 2.11</b> , Summary of findings .....	83
<b>Table 3.1</b> , Comparison between original OEP and ProAct65+ OEP.....	103
<b>Table 3.2</b> , Comparison between original FaME and ProAct65+ FaME.....	104
<b>Table 3.3</b> , Schedule of outcome measures.....	109
<b>Table 3.4</b> , PSI variables & data sources .....	128
<b>Table 3.5</b> , Bivariate analysis selection .....	137
<b>Table 4.1</b> , Falls reported in previous year, by gender and age, n (%) .....	144
<b>Table 4.2</b> , Participants' baseline falls characteristics by group, n (%).....	145
<b>Table 4.3</b> , Baseline falls-related measure means compared to normative data .....	146
<b>Table 4.4</b> , Falls Incident Rates & Rate Ratios.....	150
<b>Table 4.5</b> , Secondary falls outcomes by group (data from ProAct65+ trial report) .....	153
<b>Table 5.1</b> , PSI characteristics .....	156
<b>Table 5.2</b> , Therapist and dose effects (TDE) group baseline falls & falls risk factors compared to those excluded from the TDE group .....	159
<b>Table 5.3</b> , Comparison of those in the FaME group who did/did not return the evaluation questionnaire .....	160
<b>Table 5.4</b> , TDE group evaluation questionnaire responses used in TE analyses .....	161
<b>Table 5.5</b> , Mean falls rate, standard deviation & confidence intervals, by PSI characteristic.....	163
<b>Table 5.6</b> , Mean falls rate, standard deviation and confidence intervals, by patient evaluation of exercise .....	164
<b>Table 5.7a</b> , Baseline functional assessment score mean, standard deviation and confidence intervals, by PSI characteristic.....	166
<b>Table 5.7b</b> , Post-intervention functional assessment score mean, standard deviation and confidence intervals, by PSI characteristic.....	168
<b>Table 5.8</b> , Post-intervention functional assessment score mean, standard deviation and confidence intervals, by patient evaluation of exercise .....	169
<b>Table 5.9</b> , Correlation analyses outcomes for exercise dose (patient attendance).....	176

<b>Table 5.10</b> , Correlation analyses outcomes for PSI age & PSI attendance.....	188
<b>Table 5.11</b> , T-test scores and p values.....	189
<b>Table 5.12</b> , ANOVA scores and p values .....	191
<b>Table 5.13</b> , Mean and standard deviation for functional assessments scores between patient-rated intensity groups .....	193
<b>Table 5.14</b> , Chair rise outcomes .....	194
<b>Table 5.15</b> , Multilevel modelling; unconditional models .....	196
<b>Table 5.16</b> , Multilevel modelling; TUG .....	199
<b>Table 5.17</b> , TUG mean and SD by PSI and time-point .....	200
<b>Table 5.18</b> , Multilevel modelling; TUG .....	200
<b>Table 5.19</b> , Multilevel modelling; chair rise.....	205
<b>Table 5.20</b> , Multilevel modelling; Romberg .....	206
<b>Table 5.21</b> , Multilevel modelling; falls rate during the intervention .....	208
<b>Table 5.22</b> , Summary of multilevel modelling analyses .....	210

## List of figures

<b>Figure 2.1</b> , PRISMA flow diagram describing the search process of finding articles exploring therapist effect in exercise/physiotherapy studies. ....	65
<b>Figure 3.1</b> , Falls and falls risk factor data collection flow chart .....	118
<b>Figure 3.2</b> , Example ProAct65+ daily falls diary .....	119
<b>Figure 3.3</b> , PSI Quality Assurance Checklist .....	130
<b>Figure 4.1</b> , Percentages with at least one fall in the year preceding the study, by age & gender, compared with HSE.....	143
<b>Figure 4.2</b> , Falls flow chart: person years reflect attrition, missing data and time at risk .....	148
<b>Figure 5.2</b> , Correlation analyses for exercise dose (patient attendance) and falls rate .....	171
.....	171
<b>Figure 5.3</b> , Proportion of fallers (1+ falls) by dose of exercise and time-point .....	173
<b>Figure 5.4</b> , Correlation analyses for exercise dose (attendance) and functional outcomes.....	174
.....	174
<b>Figure 5.5</b> , Correlation analyses for PSI age and falls rate .....	177
.....	178
<b>Figure 5.6</b> , Proportion of fallers (1+ falls) by PSI age and time-point .....	179
<b>Figure 5.7</b> , Correlation analyses for PSI age and post-intervention functional outcomes .....	180
<b>Figure 5.8</b> , Correlation analyses for PSI attendance and falls rate.....	183
.....	184
<b>Figure 5.9</b> , Proportion of fallers (1+ falls) by PSI attendance and time-point.....	185
<b>Figure 5.10</b> , Correlation analyses for PSI attendance and post-intervention functional outcomes.....	186
<b>Figure 5.11</b> , Box plots for TUG by PSI .....	201

## **List of appendices**

**Appendix 1:** ProAct65+ primary outcome paper

**Appendix 2:** ProAct65+ protocol paper

**Appendix 3:** Patient evaluation

**Appendix 4:** Histograms for outcome variables

## **List of Abbreviations and Symbols**

ABC Activities-specific Balance Confidence scale  
AFRIS Attitudes to Falls-Related Interventions Scale questionnaire  
AGS American Geriatrics Society  
AMED Allied and Complementary Medicine Database  
ANOVA analysis of variance  
BGS British Geriatrics Society  
BHF British Heart Foundation  
CASP Critical Appraisal Skills Programme  
CRD Centre for Reviews and Dissemination  
CHAMPS Community Healthy Activity Model Program for Seniors  
CI confidence interval  
CINAHL Cumulative Index to Nursing and Allied Health Literature  
CONFbal Confidence in Maintaining Balance  
CONSORT Consolidated Standards of Reporting Trials  
DoH Department of Health  
FaME Falls Management Exercise  
FES Falls Efficacy Scale  
FES-I Falls Efficacy Scale-International  
FICSIT The Frailty and Injuries: Co-operative Studies of Intervention techniques  
FR Functional Reach  
GP general practitioner  
HSE Health Survey for England  
ICC intraclass correlation coefficient  
IQR inter-quartile range  
IRR incidence rate ratio  
LLT Later Life Training  
LSNS Lubben Social Network scale  
MLM multilevel modelling  
MSPSS Multidimensional Scale of Perceived Social Support  
MVPA moderate to vigorous intensity physical activity

n number of participants  
 NICE National Institute for Health and Clinical Excellence  
 NIHR National Institute for Health Research  
 OEE Outcome Expectation for Exercise  
 OEP Otago home Exercise Programme  
 ONS Office for National Statistics  
 OPQoL Older People's Quality of Life Questionnaire  
 OR odds ratio  
 PA physical activity  
 p p-value  
 PCRN Primary Care Research Network  
 PM peer mentor  
 PPA Physiological Profile Assessment  
 PRISMA Preferred Reporting Items for Systematic reviews and Meta-Analyses  
 ProFaNE Prevention of Falls Network Europe  
 PSI postural stability instructor  
 QA quality assurance  
 RCP Royal College of Physicians  
 RCT randomised controlled trial  
 RaR rate ratio  
 Romberg Modified Clinical Romberg Static Balance Test  
 RR risk ratio  
 SD standard deviation  
 SF-12 Short Form questionnaire-12 items  
 SG Sheena Gawler  
 TDE therapist and dose effect  
 TE therapist effect  
 TUG Timed-Up-and-Go  
 WHO World Health Organisation World Health Organisation  
 ZS Zoe Stevens  
 $\alpha$  alpha level  
 -2LL -2 Log-likelihood

## **Acknowledgements**

Thank you to Professor Steve Iliffe at UCL for giving me this opportunity and to Steve, Dr. Kate Walters and Professor Richard Morris for supervising me and answering my never-ending questions. I am very grateful for your patience.

Thank you to my dear Dad, John Land, who has painstakingly proof read my thesis. Thank goodness you are so much more 'detailed' than I will ever be! Thank you to my Mam, Julie Land, who has never stopped believing in me.

Thank you to Susie Dinan-Young and Dawn Skelton for your encouragement and support, and to Vicki Goodwin for being my PhD mentor.

Thank you to my partner and son who have put up with me being sat at a desk when I should have been spending time with them.

Thank you to my colleagues at UCL and a special thank you to all the ProAct65+ participants and PSIs; without you I'd have had nothing to study!

I would like to acknowledge the ProAct65+ study, which was supported by the National Institute for Health Research HTA programme, for funding this PhD.



## **Chapter 1**

### **Background**

#### **1.1 Chapter summary**

This chapter describes the context for the research described in this thesis. My research aimed to investigate the effect of the therapist and the effect of dose within a falls prevention exercise programme. The contents of this chapter therefore include the definitions of a fall, the prevalence and consequences of falling, risk factors for falling and identifying those at risk of falls. Following this, I go on to describe the evidence for exercise in the prevention of falls and the contribution of the two exercise programmes discussed in this thesis; the Otago Home Exercise Programme and the Falls Management Exercise Programme, to the evidence. The final part of this chapter includes an introduction to therapist effects, and an outline of my research aims, objectives and hypotheses. Chapter 2, the literature review, will further consider the role of therapists on patient outcomes in exercise.

It has been possible for me to study therapist and dose effects within falls prevention exercise by utilising data from the ProAct65+ Trial. My study was nested within this trial and therefore relevant detail regarding the methodology of the ProAct65+ trial is included in my thesis (Chapter 3, section 3.2).

#### **1.2 Definition of a fall**

Definitions of falls used in research typically describe the nature of a fall as being unintentional or unplanned, but although research definitions share this factor in common, different definitions have been used and this has made falls research outcome comparison difficult (Masud and Morris, 2001, Close, 2005b). As a result, there has been much interest in finding a universally adopted falls definition so that all falls research includes or excludes similar incidents during data capture, thus facilitating meaningful comparison between studies.

One such example is the inclusion or exclusion of syncopal falls (falls caused by a loss of consciousness due to a change in blood pressure) within a general definition of falls. From an epidemiological perspective, falls rates will be underestimated if syncopal falls (or indeed any other specific fall type, such as during a stroke) are excluded, however, exercise intervention studies, for example, may have excluded individuals with fall types that are not amenable to change via exercise, thus potentially making the intervention arm more effective. An example of a definition for falls amenable to exercise interventions is 'unintentionally coming to the ground or some lower level and other than as a consequence of sustaining a violent blow, loss of consciousness, sudden onset of paralysis as in stroke or epileptic seizure' (Kellogg International Work Group, 1987, Liu-Ambrose et al., 2008).

The World Health Organisation (WHO) state that the common definition for falls is 'Inadvertently coming to rest on the ground, floor or other lower level, excluding intentional change in position to rest in furniture, wall or other objects' (WHO, 2007), however, in their global report on falls in older age they are also clear to point out the difference in what older people themselves might classify as a fall compared with the view of a health professional (WHO, 2007). In line with this, Zecevic (2006) interviewed 477 community-dwelling seniors and 31 health care providers in order to define a fall. Both older adults and health care providers were concerned with the circumstances preceding the fall and also the consequences of it, whereas research definitions, the authors state, are more concerned with the fall event itself (Zecevic et al., 2006).

ProFaNE (Prevention of Falls Network Europe) established 'an unexpected event in which the participant comes to rest on the ground, floor, or lower level' as a universal consensus definition in 2005, however this has not always been adopted (Lamb et al., 2005). The simplicity of the definition is intentional to allow lay persons (study participants) to understand what should or should not be recorded as a fall. This concept is important as almost all falls captured in research are by participant self-report. It is also stated by the authors that the consensus definition is deliberately inclusive, for example a fall would not be excluded from being reported should it be considered to be an accident (for example, a collision or force knocking someone

over). Although these falls do not indicate a level of frailty that may be modified by an exercise intervention and might therefore dilute the effect of the intervention, these fallers would be randomised equally amongst the trial arms.

Hauer and colleagues (Hauer et al., 2006) conducted a systematic literature review of falls definitions. Of 90 papers that met the selection criteria, half did not provide a working definition as it was assumed the understanding by participants was implicit. The most frequently used definitions were Kellogg (Kellogg International Work Group, 1987)

- ‘unintentionally coming to the ground or some lower level and other than as a consequence of sustaining a violent blow, loss of consciousness, sudden onset of paralysis as in stroke or an epileptic seizure’

and FICSIT (The Frailty and Injuries: Co-operative Studies of Intervention techniques) (Buchner, 1993)

- ‘unintentionally coming to rest on the ground, floor or other lower level’

(8 and 9 papers, respectively) but these had frequently been modified and the remaining papers used unreferenced definitions. The lack of standardisation amongst research teams with regard to falls definition makes comparison of results less robust.

A much cited example of how the definition can influence the results is the Wolf Tai Chi study (Wolf et al., 1996) that utilised two definitions; one including and one excluding stumbles. When using the former, broader definition, results showed a significant reduction in falls in the intervention group compared to the control, whereas, the stumbles-excluded definition showed no significant reduction.

I will discuss the definition of falls used in this thesis, and the consequences of this choice in relation to what is known about falls definitions in Chapter 3 (section 3.28 FaME and OEP interventions).

### **1.3 Why study falls?**

Falls are very common in later life with potentially devastating consequences in terms of health. From a health economics perspective, falls are very costly to the NHS, and yet many falls are preventable. The following section expands upon the prevalence of falls, the consequences to those who fall and the impact of falls on the NHS.

#### **1.3.1 Prevalence of falls**

Depending on the definition of falls used, and, more importantly, the population being studied, falls rates have ranged from 0.3–1.6 per person annually in apparently-healthy community-dwellers aged 65 plus, to 0.6–3.6 per bed annually in institutionalised people aged 75 plus (Rubenstein, 2006). Older literature reports varying proportions of fallers in different age ranges of community-dwellers; 32% in those aged 75 and over (Tinetti et al., 1988), 42% in those aged 75 and over (Downton and Andrews, 1991), 34% of women aged over 65 (Lord et al, 1993), 1 in 3 women over the age of 65 and 1 in 2 men and women over the age of 85 (O’Loughlin et al., 1993). More recent publications simply state that a third (or 30%) of community-dwelling people aged 65 or over fall once or more annually and that falls in the older population are common (Gillespie et al., 2012, Nyman and Victor, 2012, Tinetti and Kumar, 2010). A rare study looking only at the ‘oldest old’ (in the UK) found that 60% of those aged 90+ years fell at least once in a year’s follow up (Fleming et al., 2008). The Lord 1993 paper still appears to be the most cited reference for falls prevalence. This study and others considered to provide reliable epidemiological falls information (Tinetti et al., 1988, Nevitt et al., 1989, Campbell et al., 1989) share in common a robust falls reporting method as well as a reliable and participant-friendly falls definition. There are many difficulties associated with obtaining accurate falls data, such as participants failing to remember falls or not wanting to report them (see section 1.6 Monitoring and reporting falls).

### **1.3.2 Consequences of falls**

These can be considered under two main sub-headings; financial cost and the consequences for individuals.

#### **1.3.2.1 Financial cost of falls**

Scuffham et al (2003) estimated that falls in older people in the UK cost the UK government £981 million annually and that almost 60% of this was contributed by the NHS. Sixty-six percent of the total costs were spent on the over 75 age group.

Englander stated that falls in the US were projected to rise from around \$20 billion in 1994 to over \$32 billion by 2020 (Englander et al., 1996). This represents more than a 50% increase in 26 years. If we apply the same increase to the UK expenditure, falls will cost £1.4 billion by 2025. In 2006 Cotter calculated that an acute hospital in Ireland was spending €10.8 million annually on falls admissions (Cotter et al., 2006). In 2010, a systematic review of cost of falls in old age was published by Heinrich et al who calculated that fall-related costs were accountable for between 0.85 and 1.5% of the total healthcare expenditure in Germany (Heinrich et al., 2010). Torgerson states that there were 86,000 hip fractures each year in the UK and that the total cost to the NHS of all falls-related fractures was in the region of £1.7 billion, based on data from 15 years ago (Torgerson and Bell-Syer, 2001). Interestingly, Grimley-Evans reported an increase in incidence of proximal femur fracture that could not be explained by the growth in older adult population size (Grimley-Evans, 1997). Speculation about the causes of this are cited as reduced dietary calcium intake, reduced physical activity and cigarette smoking. However, this study is approaching twenty years old now, so it may be that behaviour has changed. Cigarette smoking is declining (ONS, 2015), although physical inactivity is not (BHF, 2015). A lack of vitamin D from sedentary (indoor) behaviour may be a contributory factor for current bone health. A more recent study reported a decline in hip fracture incidence in England of almost 3% from 2002 to 2008, but an estimated cost of £3.6–5.6 billion in total care by 2033 for those requiring surgery for fractured neck of femur in England (White and Griffiths, 2011).

### **1.3.2.2 Consequences of falls for individuals**

The consequences of falls to the individual are far-reaching. Falls related injuries include lacerations, dislocations, fractures, head injury, haematomas and other soft tissue injuries such as sprains and ligament damage (Lord et al., 1993). Nevitt reported that only 6% of falls resulted in a major injury such as a fracture or dislocation but 55% caused a soft tissue injury (Nevitt et al., 1991). They also found that syncopal falls were more likely to cause a serious injury than non-syncopal falls. There are additionally the psychological consequences of falling including fear, anxiety and depression, and associated loss of confidence in mobilising independently, leading to an increased need for care or support. Fear of falling in community-dwelling older people ranges from 3 to 85% (Scheffer et al., 2008) although the authors of this meta-analysis state that the broad reported range is likely the result of the different measurements used (Scheffer et al., 2008). Fear of falling was originally thought to be part of the sequelae following a fall, however, it was later recognised that having experienced a fall was not a pre-requisite for fear. Fear of falling is present in 12 to 65% of older adults who have not fallen and in 29 to 92% of those who have had a previous fall (Legters, 2002). Both physical and psychological outcomes may cause a reduction in physical activity and social outings, thus potentially leaving the older person increasingly isolated and deconditioned.

Another well-documented potential consequence of falling is the 'long lie', defined as being on the floor for an hour or longer after a fall and not being able to get up (Wild et al., 1981). Inability to rise from the floor may be as a result of injury, but commonly frailer, older people who experience a long lie are not seriously injured, but lack the strength, power, flexibility and balance to perform the transfer back to their feet. Lord showed that fallers had reduced quadriceps strength and poorer static and dynamic balance compared to non-fallers of the same age and gender (Lord et al., 1992). Tinetti showed an association between age ( $\geq 80$ ), depression and poor balance and gait with inability to get up following a fall (Tinetti et al., 1993). In addition, she states that fallers who are unable to rise from the floor are more likely to have reduced ability to perform activities of daily living than fallers who can rise from the floor. Between 30% and 50% of older people requiring ambulance assistance due to a fall are not admitted

to hospital (Logan et al., 2010, Halter et al., 2000). A more recent review reported that 11 to 56% of older people were not transported to hospital following a fall (Mikolaizak et al., 2013). This gives an indication of how many UK falls are non-injurious, yet the older person needs assistance from another person to rise from the floor.

Unsurprisingly fear is more prevalent in those older people who are unable to rise from the floor. The long lie itself can have other serious physical consequences including pressure sores, hypothermia, dehydration, bronchopneumonia, kidney failure and death (Masud and Morris, 2001). In line with this, falls services often offer 'education' sessions to advise fallers how to offset some of the consequences of a long lie, should they find themselves unable to transfer up from the floor.

Lord published data that show outcomes for fallers compared to age and sex matched non-fallers are poor (Lord et al., 1992). Within a year, 27% of fallers had either had 3 or more additional falls, been readmitted to hospital, moved to a nursing home or had died. Falls are responsible for two-thirds of all deaths caused by injury and half of all older people admitted to hospital following a fall will have died a year later (Rubinstein, 2006). Injury resulting from falls is the leading cause of death in people aged over 75 in the UK (DoH, 2010). Those older people who experience recurrent falls have increased susceptibility to hospitalisation, institutionalisation and increased mortality (DoH, 2009).

#### **1.4 Risk factors for falling**

It is thought that there are over 400 risk factors for falls (Oliver et al., 2000). As with coronary heart disease, an accumulation of several risk factors is more likely to cause a fall than a single risk factor. In line with this, Nevitt found that 69% of subjects who had experienced two or more falls in the previous year demonstrated four or more risk factors, whereas only 10% of subjects who had experienced two or more falls demonstrated none or one risk factor (Nevitt et al., 1989). Tinetti published similar findings: the proportion of recurrent fallers increased from 0% of those with 0 to 3 risk factors, to 31% of those with 4 to 6 risk factors and 100% of those with 7 or more risk

factors (Tinetti et al., 1986). The multi-factorial nature of falls and the large number of recognised risk factors means that no two falls are the same.

Risk factors can be sub-divided into intrinsic and extrinsic categories (Todd and Skelton, 2004). Intrinsic factors pertain to the individual whereas extrinsic ones are environmental, such as poorly maintained pavements, slippery indoor flooring, clutter within the home etc. Close (2005) groups intrinsic risks; demographic, mobility/balance, gait, sensori-motor, medical and medications, and rates each risk according to the importance of its association with falls (on the basis of the strength of the evidence). Risk factors with the highest rating include advanced age, history of falls, impaired gait, poor vision, reduced muscle strength, impaired reaction time, stroke, Parkinson's, impaired cognition and multiple medications (Close, 2005b). Female gender is less closely associated with increased risk as it is usually but not always a finding in previous research (Close, 2005b). Impaired balance, mobility and gait are very strongly associated with falls indicated by the fact that all risk factors in these domains have the highest rating (Close, 2005b). A more recent meta-analysis found that having fallen is still the strongest predictor of a future fall (Deandrea et al., 2010).

Campbell and colleagues noted that risk factors associated with falls differed for men and women (Campbell et al., 1988). Increased susceptibility to falls in men was associated with reduced physical activity, stroke, knee arthritis, impaired gait and increased body sway. In women, however, polypharmacy, psychotropic drugs, drugs associated with postural hypotension, low standing blood pressure and muscle weakness were implicated in falls. Another study found that there were gender differences in risk factors for injurious falls (Koski et al., 1996). Women were more susceptible to falls associated with culprit medications, such as benzodiazepines, whereas men's injurious falls were associated with quadriceps dysfunction and gait problems (Koski et al., 1996). More recent data from the USA show that men experience fewer falls than women, but are more likely to have a fatal fall. This is attributed to men having more co-morbidities than age-matched women (Stevens et



al., 2007). In 2012 the same author reported that women fallers were significantly more likely to see a healthcare professional after a fall (Stevens et al., 2012).

Risk factors vary for different populations, for example Lord published data suggesting that nursing home and intermediate-care residents' risk factors were different (and predictable) based on their ability to rise from the chair and stand unaided (Lord et al., 2003).

Risk factors need also to be considered in terms of their potential for remedy. Risks are therefore sometimes referred to as being either modifiable or non-modifiable with exercise. Advanced chronological age is clearly non-modifiable, whereas poor muscle strength and poor balance are modifiable. Other non-modifiable risk factors are impaired vision and hearing, having chronic health problems, foot deformities and recent discharge from hospital (Lord et al., 2007). Exercise may also have a less direct effect on falls risk factors; for example, it could reduce hypertension thus potentially reducing the dose or number of anti-hypertensive medications. If this in turn reduces the symptoms of postural hypotension, falls may be reduced.

The acronym DAME has been used as an aide memoir for categorising risks; Drugs and alcohol, Age-related physiological decline, Medical conditions and Environmental risks (Dinan and Skelton, 2012). This clearly includes extrinsic as well as intrinsic risks. Environmental hazards, it could be argued, are often unavoidable and with this in mind, anyone can trip over something and have an accident. However, it is the intrinsic ability to cope with uneven surfaces, perturbation and trip hazards, for example, that identifies fallers from non-fallers. Non-fallers are more likely to prevent a trip becoming a fall by using a balance correcting strategy, such as taking a compensatory step. Several studies have compared the reflex actions of younger subjects versus older subjects or fallers with non-fallers when their postural stability was perturbed. Luchies found that in response to a backwards pull from the waist, younger subjects took one backwards compensatory step in order to regain their balance, whereas older subjects took several, shorter steps suggesting that the first step was insufficient to regain their postural stability (Luchies et al., 1994). Another perturbation study

comparing frequent fallers with non-fallers found that displacing balance in an anterior-posterior plane caused frequent fallers to take more lateral steps to recover their balance suggesting that those with impaired balance have more difficulty stabilising the body laterally when the displacing force is in the anterior-posterior plane (Rogers et al., 2003).

There is some interesting cross-over between environmental and intrinsic factors; for example, cold weather conditions will negatively impact muscle strength and power as muscle is approximately 50% less able to contract at a temperature of 15°C (Skelton et al., 1992) but if the person has underlying poor muscle power, the combination of that and the effect of cold weather on muscle might cause a fall. In a more recent study investigating the effect of cold temperature on muscle function, exposure to a room temperature of 15°C had a significant reduction on leg extensor power, as well as some key functional indicators; sit-to-stand performance and gait speed (Lindemann et al., 2014). Another consideration is at the interface between intrinsic and environmental risks; for example, footwear, eyeglasses and medications.

Risks have also been considered as psychological versus physical. Depression and dementia are known risk factors, as is fear of falling. A group of community-dwelling older people with cognitive impairment had a falls incidence rate of 67% (Tinetti et al., 1988) and another study reported that as many as 85% of a group of subjects with dementia fell in a year (van Dijk et al., 1993). Fear is predictive of falls in its own right i.e. independently of additional physical risks. This is thought to be accounted for by the fact that people tend to alter the way they move, thus altering their normal functional movement patterns, when they are fearful. A gait study published in 2005 found that older people with fear of falling (but not necessarily a history of falls) had a shorter stride length, greater stride width, a slower gait speed, and spent longer in double stance than those without fear of falling (Chamberlin et al., 2005).

### 1.5 Identifying those at risk

One of the strongest risk factors for a fall is a previous fall, and national guidance states that all older people in regular contact with healthcare professionals should be asked about previous falls at least once a year (NICE, 2013).

The Timed Up and Go Test (TUG) is recommended by the American Geriatrics Society/British Geriatrics Society and is widely recognised as a simple, quick assessment for identifying those at risk of falls. The TUG validation paper was published in 1991 (Podsiadlo and Richardson, 1991). They recruited older adults referred to a Geriatric Day Hospital and showed TUG's ability to predict the patient's ability to safely mobilise outdoors showing that poor TUG scores were associated with poor independent mobility. The authors also reported that taking less than 30 seconds to perform TUG was associated with functional independence. A number of subsequent TUG studies have focused on TUG's use as a tool to identify fallers from non-fallers within the older population. Shumway-Cook recruited American, community dwelling older adults aged 65 or over and found a 90% correct prediction of fallers (having reported 2 or more falls within the last 6 months) for those taking 13.5 seconds or more to perform TUG (Shumway-Cook et al., 2000). However, Rose (2002) found that 10 seconds could be used as the reliable cut-off to identify fallers, and Whitney, who recruited patients aged 63-95 who had been referred to a falls clinic in a London Hospital, reported that a 15 second cut-off point to identify those at high falls risk gave maximal sensitivity and acceptable specificity (Whitney et al., 2005). In the Whitney study, those at high risk were identified in the first instance by Physiological Profile Assessment (PPA) - scoring more than 2 – then TUG was used to identify the same cohort (those scoring >2 on PPA).

Physiological Profile Assessment (PPA) allows more optimal tailoring of interventions on the basis of a more detailed assessment of physiological impairments. The PPA data are divided into domains: vision, proprioception, muscle strength, reaction time and postural stability. Whitney (2005) found a significant association between TUG and PPA scores (and a TUG cut-off point of 15 seconds to identify those at high risk of falls).

PPA is expensive to administer and is therefore not widely used, and considering that other, cheaper tools can accurately predict those likely to have future falls, its use has been recognised as being more useful for tailoring interventions than identifying fallers. There is some speculation, however, as to whether TUG alone is useful for discriminating fallers from non-fallers (Schoene et al., 2013) in a high-functioning, healthy older adult population.

The Functional Reach test, although not typically used as a stand-alone assessment of falls risk, has also been found to have some predictive usefulness. Duncan and colleagues evaluated the Functional Reach (FR) assessment in American adults aged 21-87 years and found it was a reliable and highly reproducible measure of balance (Duncan et al., 1990). In men aged 70-87, the mean FR was 33.43cm and in women of the same age it was 26.59cm. In 1992 the same research team investigated FR as a potential measure of falls risk in community-dwelling men aged 70 or older. They found that a reach of 6 inches (15.24cm) or less predicted having 2 or more falls in the 6 month follow up period. The authors concluded that FR can therefore be used to identify recurrent fallers. One study (Lin et al., 2004) compared the accuracy of FR and TUG as tests for identifying fallers in community-dwelling older adults and reported receiver operating characteristic (ROC) area under the curve (AUC) of 0.51 and 0.61, respectively, suggesting that TUG is the more accurate assessment.

Many studies and indeed falls services (Gates et al., 2008) use the Falls Risk Assessment Tool (FRAT) to identify those at high risk of falls. FRAT was developed by Nandy and is intended for use in primary care or the community (Nandy et al., 2004). It comprises five questions that take into account several of the falls risk factors identified by Close (2005b); namely, medical risks (a diagnosis of stroke or Parkinson's), medications (4 or more), previous fall(s) in the last year, lower limb muscle weakness (unable to rise from a chair without using the upper limb) and poor self-reported balance. The FRAT validity study (Nandy et al., 2004) recruited 1000 people aged 65 and over and reported that 3 or more risk factors predicted a fall in the

next 6 months with a positive predictive value of 0.57 (95% CI 0.43-0.69) i.e. 57% of those identified as high risk reported a fall in the following 6 months.

As fear of falling is an independent risk factor for falls, there are published tools to identify those with fear. In a systematic review, Scheffer and colleagues identified ten tools for measuring fear of falling and categorised them according to the construct used to measure fear; fall-related self-efficacy or fear of falling (Scheffer et al., 2008). The most prevalent fear of falling measure in the included literature was the single question "Are you afraid of falling?". The most frequently used self-efficacy measures were the Falls Efficacy Scale (FES) and the Activities-specific Balance and Confidence Scale (ABC). The short Falls Efficacy Scale-International (Short FES-I) contains 7 domains (Kempen et al., 2008) each with a possible score of 1 through 4 (1=not at all concerned, 4=very concerned) (Yardley et al., 2005). A higher total score indicates poorer self-efficacy, with a maximum possible total score of 28. A published cut off point of 11 differentiates between low and high concern about falls for a range of activities of daily living (Delbaere et al., 2010). The ABC is a measure of confidence to maintain balance and may be more suitable than FES for assessing loss of confidence in more functionally able older adults (Powell & Myers, 1995). The Confidence in Maintaining Balance (CONFbal) questionnaire, another measure of balance confidence, was developed specifically for frailer, older adults (Simpson et al., 1998). CONFbal contains 10 questions regarding everyday activities (such as getting up from a chair and walking). Each question has three possible responses; confident, slightly confident and not confident, which are awarded a score of 1, 2 and 3, respectively (Simpson et al., 2009). A higher total score indicates poorer confidence, with a maximum possible total score of 30. There are no published cut-off points to identify those who are not confident they will maintain their balance during the specified everyday tasks.

The best measure for identifying those at risk depends on the setting. Scott and colleagues reviewed 38 falls prediction tools and found that most tools had not been assessed for validity and reliability more than once, nor in more than one setting, suggesting therefore that no single tool can be recommended for all settings (Scott et al., 2007). As well as this, the tool should be selected on the basis of the purpose for

which it is being administered. In general practice, for example, screening for high-risk patients via asking them if they have fallen is time-efficient and, as a previous fall is so highly predictive of future falls, effective secondary prevention via appropriate onward referral is potentially effective, without the use of any validated 'tool'. In physiotherapy services, however, tool selection may be determined by the tool's ability to identify remediable risk factors, thus informing intervention choice. In the research setting, study population can be more effectively characterised if the number of falls in the previous year are recorded. One reason for this might be to identify frequent fallers (those who have fallen three or more times in the previous year) who have different characteristics to those who fall less frequently (Masud and Morris, 2001, Koski et al., 1996). Problems associated with reporting falls in the research setting are discussed in the following section (1.6 Monitoring and reporting falls).

### **1.6 Monitoring and reporting falls**

In intervention studies (or indeed falls services) aiming to reduce falls there is clearly a need to collect data about the number and severity of falls. Falls data collection commonly occurs during interventions to compare falls rates between intervention and control arms, but some studies have collected falls data for up to a year prior to intervention start in order to see if falls are reduced in the same arm during (and after) the intervention, in comparison to before it. In either case, accurate methods of collecting falls data need to be employed and there have been several trials comparing the effectiveness of retrospective and prospective reporting methods. It is now agreed that the prospective reporting eliminates under-reporting inaccuracy caused by subjects forgetting their falls and in some cases over-reporting caused by 'telescoping' (subjects forget when the fall occurred and 'telescope' it into a reporting period in which it did not occur) (MacKenzie et al., 2006). The percentage of under-reporting by retrospective report ranges from 13% (Cummings et al., 1988) to 25% (Hannan et al., 2010). Cummings et al additionally noted that contrary to expectations, recall was better over a 12-month period than in shorter time-frames (3 or 6 months) suggesting that some older people (over 60) did not forget a fall completely, they just forgot when it occurred. As might be expected, lower scores on the Mini-Mental State examination

were associated with poor falls recall. The Hannan reliability study (2010) compared daily falls reporting on diaries (returned monthly) with 3-month recall by telephone call. Both methods accurately identified non-fallers but the retrospective recall missed up to 25% of falls.

In 2012 Perry published data from the ProAct65+ trial attempting to predict successful falls report completion and return on the basis of education level, first language spoken and baseline falls risk (Perry et al., 2012). Interestingly, fallers (identified by FRAT) were less likely to return their diaries and, as expected, native English speakers were more likely to complete their diaries correctly. This research highlights the fact that under-reporting is possible even when utilising the recommended prospective method of reporting and that missing data are not likely to be randomly missing. This is of most importance in falls studies recruiting individuals irrespective of their baseline falls status and in culturally diverse populations.

There has been some speculation as to whether forgotten falls are likely to be the least serious in terms of injury or other consequence. In 2012, Delbaere and colleagues reported that in community-dwelling older people, better memory was correlated with falls (Delbaere et al., 2012). They suggest that the higher chance of an older person with better memory being classified as a faller is a reporting bias. They also note that if injuries resulting from falls are part of the definition, those with better memories do not appear to fall more. This suggests that injurious falls are remembered equally by those with better and poorer memory alike. However, another study (Mackenzie et al., 2006) found that self-report of injuries was poorer than self-report of falls.

Another debated falls reporting issue has been that of raised awareness of falls in intervention studies. Mackenzie found that recalled falls were under-reported by 13% in comparison to falls that were prospectively reported, and that recall was more accurate in the intervention group compared to the control group (Mackenzie et al.,

2006). Without their subsequent adjustment for study arm, the intervention would appear to have increased falls compared to the control.

The FaME study (Skelton et al., 2005) collected falls data for at least 36 weeks prior to intervention start. Subjects completed daily falls diaries that were returned fortnightly. This data allowed the authors to calculate a mean falls rate at baseline in both trial arms. As the study was selecting frequent fallers, the prospective monitoring of falls prior to the start of the intervention was needed to accurately identify those who had 3 or more falls in the year preceding the study. This study is unusual in that it collected falls data pre-intervention. These data also allowed the research team to adjust for baseline falls in the between group analysis to calculate that the number of falls in the intervention group fell progressively from baseline, whereas the falls rate in the control group did not fall.

In 2005 ProFaNE recommended that fall rates should be calculated from prospective monthly diaries (Lamb et al., 2005), however, in their systematic review of methods of measuring falls in trials, Hauer found that only 41% of fall prevention intervention studies used prospective data collection (Hauer et al., 2006). Since this was published only one year after the ProFaNE guidelines were published, there is still every possibility that future studies will adopt the recommended method, prospective data collection with frequent returns of diary, even though this gold-standard method of falls reporting is burdensome for participants. The frequency of return of prospective diaries has ranged from 1 week to 6 months (Hauer et al., 2006) but there does not appear to be an optimal frequency. The burden of falls diaries on participants can result in missing data. Forty-eight percent of monthly diaries were returned in one study (Hogan et al., 2001) and another (Perry et al., 2012) reported that 40% of the subjects returned all diaries.

Other studies have considered alternative reporting methods such as calling a falls 'hotline' and using surveillance data, such as nursing home or hospital fall records (Kerse et al., 2004). van der Marck and colleagues reported that 100% of no fall reports and 78% of fall reports via an automated falls telephone were found to be correct.



They point out that the falls telephone can reduce the volume of time-consuming personal follow-up calls in particular to non-fallers. This method would therefore be especially helpful in large, long studies and with groups whose falls risk is low (van der Marck et al., 2011).

### **1.7 Outcome measures**

Since outcome measures should be selected to reflect the study hypothesis, falls prevention intervention studies should clearly collect numbers of falls (via prospective reporting) as the primary outcome, however, as discussed in the previous section, it is difficult to get accurate data that are free from reporting bias.

In addition to falls rate, many studies have collected data about injuries resulting from falls. Collecting this information may allow fall severity comparison between intervention and control, and clearly, should fall severity be reduced but falls rate not be significantly reduced, this would still be a positive outcome for the individual. In the case of the Campbell Otago Home Exercise studies (Campbell et al., 1997, Campbell et al., 1999a, Campbell et al., 1999b, Robertson et al., 2001a, Robertson et al., 2001b), falls were reduced by 35% overall and injurious falls were reduced by the same percentage.

Another strategy is to assess changes in remedial falls risk factors. The outcome measures utilised should be matched to the intervention; for example, an exercise intervention aimed at reducing falls via improving lower limb muscle strength and balance should employ sensitive and validated tools for assessing muscle strength and balance. The benefit of this approach is avoiding the participant and researcher burden associated with collecting falls data, but the accompanying limitation is the lack of evidence that the intervention reduces falls.

Chair rise assessment is a useful test of functional leg power. The number of chair rises performed in a set time (typically 30 seconds) is appropriate for the general older population, otherwise, with frailer populations, the time taken to perform a set

number stands (typically 5) is employed. Jones and colleagues published normative Functional Fitness scores (including the 30-second chair rise) in 7000 community-dwelling American adults aged 60-94 (Jones et al., 1999). The mean number of chair stands performed in 30 seconds typically declines from 15 to 13 in women, and 16 to 14 in men, over three decades.

The physiotherapy guidelines for the management of fallers (Goodwin and Briggs, 2012) recommend the use of reliable and sensitive outcome measures, in particular, to assess balance. Tinetti's (1986) and Berg's (1989) balance assessments are commonly used in physiotherapy practice. They are validated and focus on gait and functional balance (Tinetti) or functional balance alone (Berg). The American Geriatrics Society and British Geriatrics Society (AGS/BGS) clinical practice guidelines for falls prevention recommend the assessment of gait and balance using either the Get Up and Go test (Mathias et al., 1986), the Timed Up and Go test (TUG) (Podsiadlo and Richardson, 1991), the Berg Balance Scale (1989) or the Performance-Oriented Mobility Assessment (Tinetti, 1986, Tinetti et al., 1988). Both Berg and Tinetti tools are rather slow to administer (taking up to 20 minutes) as they are thorough, however, this is appropriate for a physiotherapy assessment of balance. In addition, the Tinetti assessment requires the assessor to be knowledgeable about the phases of gait, which may render it inappropriate for use other than by a specialist. In falls exercise research a shorter outcome measure may be more appropriate, given that the number of outcome measures used is often greater than in clinical practice. TUG test, for example, is quicker to administer than either Berg or Tinetti, however, the AGS/BGS report highlights the lack of validation for all the aforementioned tools (Get Up and Go test, TUG test, Berg Balance Scale & Tinetti) at identifying future fallers in a population of general older adults. A combination of assessments may therefore be the best available approach in the research setting.

## 1.8 Falls prevention exercise research

There is already robust evidence to support the use of exercise in falls prevention and rehabilitation following falls, with as many as 42% of falls being prevented by a “well-designed” exercise intervention (Sherrington et al., 2008).

The updated Cochrane Systematic Review of Interventions for preventing falls in older people living in the community, states that both group and home based exercise reduce falls rate (Rate Ratios 0.71, 95% CI 0.63 to 0.82 and 0.68, 95% CI 0.58 to 0.80, respectively) and risk of falling (Risk Ratios 0.85, 95% CI 0.76 to 0.96 and 0.78, 95% CI 0.64 to 0.94, respectively) (Gillespie et al., 2012). Tai chi was found to significantly reduce falls risk (Risk Ratio 0.71, 95% CI 0.57 to 0.87), but not the rate of falls. The review also includes available evidence for multifactorial interventions, vitamin D, home safety assessment, treatment of vision problems, pacemakers, withdrawal of psychotropic medication, cognitive behavioural interventions and fall prevention education alone. Group exercise, home exercise and home safety are the only three interventions that reduce both falls rate and risk of falls. The authors also reported sub-group analyses for baseline falls risk i.e. the effectiveness of the interventions in those with low baseline falls risk and separately in those with high baseline risk (previous falls history or risk factor(s) for falls). Group exercise was effective in both high and low risk sub-groups for falls rate and risk of falling. Tai chi, however, was more effective for those without previous falls or pre-existing risk factors for falls i.e. in primary falls prevention. Interestingly, of the seven exercise intervention trials that compared different types of exercise, one study (Kemmler et al., 2010) investigated higher versus lower intensity multi-component group exercise and found that falls rate was 40% lower and the number of fallers was 46% lower in the high intensity group. When exercise interventions were investigated according to the exercise modality (for example, balance, strength, endurance) no single-modality intervention was successful in reducing either falls rate or the number of fallers, with the exception of those in the ‘balance, gait and functional training’ category which significantly reduced falls but not

fallers. This reinforces the message that balance training is arguably the most important exercise type when aiming to reduce falls.

### **1.8.1 Falls prevention exercise prescription**

In 2011 Sherrington and colleagues published their updated meta-analysis of falls prevention exercise trials and accompanying best practice recommendations (Sherrington et al., 2011). Interventions that contain balance retraining exercises, a total dose of at least 50 hours (2 hours a week for 6 months) and no walking training, seem to be the most effective (38% reduction in falls). Balance exercise that is moderately or highly challenging is indicated. Trials that did not recruit individuals on the basis of falls risk (unselected individuals) demonstrated a greater reduction in falls; 27% compared to 10% in a high risk population. This is counterintuitive given that those at high risk have more falls and therefore there are more falls to prevent. The authors of the meta-analysis do not suggest why exercise used in the primary prevention of falls appears to be more effective than in secondary prevention. Perhaps the apparently low risk general older adult population have a higher falls rate over the trials' follow up periods than would be predicted by their baseline falls status. The best practice guidelines (Sherrington et al., 2011) suggest that:

- 1) High to moderate challenge balance exercise should be achieved by reducing the base of support, moving the centre of gravity and by reducing the need for upper limb support (holding on),
- 2) Exercise should be of a high dose (more than 50 hours),
- 3) Falls prevention exercise should be offered to the general, older population as well as selected individuals who are at high risk of falls,
- 4) Falls prevention exercise can be offered as group-based or home-based,
- 5) Balance training may be offered in combination with strength training,
- 6) Walking is used with caution for high risk individuals and safety to perform walking should be assessed.

### **1.8.2 Walking**

There is some concern that increasing physical activity, particularly brisk walking, may increase exposure to risk of falls (Sherrington et al., 2011). This first emerged in a trial

of brisk walking versus upper limb exercise to improve bone mineral density in women with previous fracture; Ebrahim and colleagues reported a significant increase in risk of falls in the self-paced brisk walking group compared with the placebo (Ebrahim et al., 1997). Faber and colleagues also showed an increase in falls rate in institutionalised frail and pre-frail older adults randomised to a functional walking group compared with the control group (Faber et al., 2006). A more recent longitudinal study found a significant association between habitual walking and falls in higher risk community-dwelling individuals but in those at lower risk, habitual walkers had significantly fewer falls than non-walkers (Okubo et al., 2015).

### **1.9 Clinical practice documents**

The updated NICE clinical guideline on falls (NICE, 2013) recommends individually prescribed muscle-strengthening and balance training for community dwelling over 65s who have a history of falls or who have a balance and gait deficit. In addition, it suggests that there is no evidence for brisk walking to reduce falls risk, and that untargeted group exercise should not be encouraged. This is surprising given that there is robust evidence for exercise in the primary prevention of falls. Physiotherapy guidelines for the management of older people at risk of falling (Goodwin and Briggs, 2012) highlight similar messages; exercise programmes should be of a high dose (50 hours over 6 months), have a balance component that is highly challenging, and that exercise can be delivered as a single intervention or as part of a multifactorial one. Furthermore, the authors suggest that strength training is also delivered, that balance training needs to be sufficiently dynamic and that reliable, valid and sensitive outcome measures are utilised to assess balance. Physiotherapy approaches stress the importance of re-training (or maintaining) the person's ability to rise from the floor to help prevent the possible complications of a long lie following a fall (Goodwin and Briggs, 2012).

The American Geriatrics Society (AGS)/British Geriatrics Society (BGS) Clinical Practice Guideline: Prevention of Falls in Older Persons (2011) states that exercise including strength, balance, gait, and coordination training, should be offered as part of a

multifactorial intervention to prevent falls or as a single intervention. Furthermore, it recommends that exercise programming should consider the functional ability and health of the older person and be prescribed by a qualified health or fitness professional. As well as being tailored to the individual, the falls prevention exercise should be reviewed and be progressive.

### **1.10 Falls exercise provision**

In 2012 the Royal College of Physicians (RCP) published findings from a questionnaire completed by older people following participation in exercise delivered by falls prevention services nationally. The questions were designed to assess the quality of older people's experiences of exercise and to map provision against published falls prevention guidelines. The report states that the evidence base for exercise indicates a prescription involving correct type (targeted resistance and dynamic balance), frequency (a minimum total dose of 50 hours, exercise occurring ideally 3 times per week), duration (15 to 52 weeks) and intensity (sufficiently challenging and progressive). They found that 81% of patients who attended a class-based intervention, were only offered 12 sessions or less, and only about half of patients believed their programme was progressive. The report also highlights that only 29% used ankle weights in order to achieve the evidence base regarding exercise type (resistance). It is interesting that a further 34% of patients used an exercise band which also provides resistance. Positive outcomes listed are that 76% found the exercise beneficial and that over 90% were satisfied or very satisfied with their programme (RCP, 2012).

Both FaME and OEP programmes are recommended in the Department of Health Prevention Package (DoH, 2009). The key messages in this resource entitled Falls and fractures; Exercise Training to Prevent Falls are 1) dynamic balance, strength and floor-based exercises should be included, 2) the exercise should be individually tailored and 3) home or group-based exercise is effective. These were the two exercise

programmes that were used in the trial that this thesis is based on. The following sections describe the evidence for their use.

#### **1.10.1 Falls interventions: Otago Home Exercise Programme**

The Otago Exercise Programme (OEP) was developed in New Zealand approaching 20 years ago, and is designed to reduce falls by improving balance and muscle strength. Patients perform 30 minutes of exercises three times per week in their home, and are supported with home visits and telephone calls. Five trials of the OEP (Campbell et al., 1997, Campbell et al., 1999a, Campbell et al., 1999b, Robertson et al., 2001a, Robertson et al., 2001b, Campbell et al., 2005) were conducted by the original research team in New Zealand, plus several studies since then, conducted by other teams (Brown et al., 2009, Liu-Ambrose et al., 2008, Kyrdalen et al., 2014, Thomas et al., 2010).

The first trial (Campbell et al., 1997, Campbell et al., 1999a) recruited 233 community dwelling women aged 80 and older from general practices in Dunedin, New Zealand, and randomised them in to either a year's individually tailored home exercise programme (Otago) or the control. Inclusion criteria included ability to move around in their own home and not receiving physiotherapy. Those subjects randomised into the exercise arm were visited at home by the physiotherapist, initially to prescribe a set of moderate intensity exercises, then a further 3 times in the first 2 months. Each visit lasted approximately an hour. Participants were asked to complete their 30 minute programme 3 times per week and to walk outside the home at least 3 times per week. To assist motivation and therefore compliance, the physiotherapist telephoned each participant regularly. Those in the control group received an equal number of social visits and telephone calls. Follow up assessments were completed at 6 months but subjects were encouraged to continue exercising for 12 months and falls and frequency of exercise were monitored for the entire 12-month intervention.

The 6-month assessment results showed a significant improvement in the 4-test balance score and the chair stand test in the exercise group compared to the control, but there was no difference between groups for the other functional assessments

(Campbell et al., 1997). At 12 months, the mean rate of falls per year was lower in the exercise group than the control, and injuries following falls were significantly less prevalent in the exercise group. Despite the increased risk of falls in the intervention arm due to increased activity, the time to first fall was similar in both groups indicating programme safety (Campbell et al., 1997).

The second paper was published two years later in 1999 (Campbell et al., 1999a) on longer term follow-up of participants in the first OEP trial. Since there is no reason to think that falls rate reduction benefits of exercise will persist unless the person continues to exercise, the authors not only looked at whether the programme was successful at reducing falls over the extended period but also at the characteristics of those who continued to perform the OEP. There were no further home visits in the second year of the programme, but the physiotherapist remained in telephone contact. The difference in falls rate between the exercise and control groups remained significant into the second year. Injuries resulting from falls also remained lower in the exercise arm. Women with a higher recorded level of physical activity and those taking fewer medications at baseline were more likely to continue with the programme in to the second year, as well as those whose falls efficacy scale score was higher after 1 year (i.e. those who perceived that the programme was making them less likely to fall after the first year were more likely to keep performing the exercises for a second year). This also suggests that those who are more fearful of falling are less likely to participate in/adhere to the programme. The authors suggest that improving compliance in this group might be achieved by offering group exercise, hip protector pads or family involvement in the intervention. At the end of the second year, 44% of participants were still doing the exercises at least three times a week. Those with a history of falls prior to starting the programme were more likely to keep going for the 2 years. The authors recommend that 6 monthly home visits are maintained (into the second year and beyond) to keep people motivated and to allow for any necessary programme modifications.

In the same year Campbell and colleagues published a paper investigating the withdrawal of psychotropic medication and the OEP (Campbell et al., 1999b). The



recruited mixed gender group were aged 65 and over and it was found that although the medication withdrawal group did experience a reduction in falls rate compared to those who did not reduce their medication, the exercisers did not demonstrate a significantly reduced falls rate compared to those who did not exercise. This perhaps provides additional evidence that the younger old population are not sufficiently at risk of falls to benefit from the programme.

A further two trials were published by the same research team two years later (Robertson et al., 2001a, Robertson et al., 2001b), both with primary objectives of assessing the effectiveness of using trained nurses to administer the OEP (rather than the research physiotherapist used in the previous two trials) and investigating the use of the programme in healthcare settings (rather than a tightly controlled research setting). In line with the latter objective, a mixed gender population was recruited in both trials. The lower age limit was lowered from 80 to 75 years in the first 2001 trial (Robertson et al., 2001a) plus an economic evaluation was completed for both. The training for the nurses was a week in duration and delivered by the research physiotherapist. Once the nurses were supporting participants to perform the programme, the same physiotherapist conducted site visits and made telephone calls to ensure quality. The programme duration, intensity, progression and visit/telephone call schedule was as the previous two studies. The use of trained nurses proved successful as the reduction in falls in the exercise group compared to the control was 46% in the first 2001 trial (Robertson et al., 2001a). However, this reduction in falls was only seen in those aged 80 and over. In the second trial falls were reduced by 30% in the exercise centres compared with the control centres.

Given this evidence for the effective use of trained nurses to deliver the OEP, it is noteworthy that this approach does not appear to have been widely implemented in the UK. However, a national training organisation, Later Life Training, have developed, in partnership with the OEP authors, an Otago Leader training course which is open to those who are neither qualified allied health professionals nor qualified exercise professionals. This course attracts a small number of nurses and many more non-qualified physiotherapy/rehabilitation assistants, indicating the adoption in the UK of

an OEP implementation model that does not rely only on physiotherapists. It could be argued that this is the translation of the 2001 OEP evidence into practice.

The cost of implementing the OEP was calculated as \$NZ432 (£138) per person for one year in the first 2001 trial (Robertson et al., 2001a) and \$NZ418 (£121) in the second (Robertson et al., 2001b). These calculations included equipment costs, recruitment costs, programme prescription costs (nurse time and travel and telephone calls, for example) as well as quality control costs (physiotherapist time and travel). The economic analyses used the control group as a comparator in terms of healthcare costs related to falls. Although the number of injuries in the exercise groups compared to the control groups in both studies was not significantly different, there were 5 fall related hospital admissions in the control group and none in the exercise arm in Robertson et al 2001a. The cost of these was calculated as \$NZ47,818 (£21,159) and this sum is therefore the amount of money that was averted in the exercise group. Using both the OEP implementation costs and the healthcare costs avoided by the OEP, a cost saving was seen in those aged 80 years or over. There was no difference between groups in hospital costs resulting from falls in the second 2001 trial.

The final study by the original research team (Campbell et al., 2005) recruited community dwelling women and men aged 75 and over with severe visual impairment and randomised them into one of four study arms; the OEP with vitamin D supplementation, a home safety programme, OEP and home safety, or social visits. The rationale for using the OEP and/or home safety for a selected older population was to try to provide some answers regarding which falls prevention programmes might be effective with specific older populations who fall (rather than implementing costly multifactorial interventions at a general older population level). Exclusion criteria were as previous OEP studies as well as those whose visual acuity was greater than a specified level.

The home safety assessment and modification programme was implemented by an occupational therapist (OT) and was designed specifically for severe visual impairment. The OT recommended home modifications/equipment installation (stair rails, grab

bars etc.) and/or behavioural modification (removing clutter) for each person and adherence was monitored at six months by telephone interview. The OEP was implemented by a physiotherapist and followed the same model as the previous studies. The only adaptation for poor vision was the supplementation of vitamin D (as those with severe visual impairment were deemed less likely to mobilise outside the home than older people without visual impairment). Those in the control group received 2 home visits each lasting an hour.

The results showed a 41% reduction in falls in the home safety programme group compared to those who did not receive this intervention and a 15% increase in falls in the exercise group compared to those who did not receive the OEP. It may have been concluded that those exercising were increasing their activity levels and therefore, opportunity to fall, however, the authors performed a separate analysis investigating falls rate by level of adherence to the OEP and found that those achieving 3 or more sessions per week had a 77% lower rate of falls than those exercising less than once per week. It was concluded that the OEP is ineffective at reducing falls in those older people with severe visual impairment and that this is likely to be as a result of poor adherence. Another conclusion is that home safety is an important part of falls prevention in the visually impaired.

Since this study a number of other research teams have used the OEP to investigate other outcomes such as cognition. In 2009 Brown and colleagues published their findings on the effect of group OEP (and group stretching/relaxation, both compared to no exercise) on cognitive performance and mood in older adults living in retirement homes (Brown et al., 2009). The hour long group sessions were twice-weekly for 6 months. Three domains of cognitive performance (fluid intelligence, memory and executive functioning) were assessed at baseline and immediately post-intervention, as well as the Geriatric Depression Scale (GDS) to ascertain any changes in mood. Whilst both the group exercise sessions improved specific aspects of mood, only the OEP exercise cohort showed improvement in the fluid intelligence domain of cognitive performance. Fluid intelligence is the ability to think creatively and problem-solve. There doesn't seem to be a logical reason why this would be improved via the

performance of one type of exercise intervention compared with another, plus, of the 3 domains studied, only 1 domain changed, so perhaps this was a chance finding.

The OEP was used by Liu-Ambrose and colleagues (2008) to investigate the contribution of executive functioning to falls. They recruited community dwelling subjects aged 70 and over with a recent history of falls to perform the OEP for 6 months and, along with physiological falls risk and prospective falls monitoring for a year's follow up, they assessed three central executive functions; set shifting, updating and response inhibition. Set shifting is a form of multi-tasking, updating is working memory and response inhibition involves purposefully inhibiting automatic responses. At 6 months the intervention arm showed a significant improvement in one aspect of executive functioning (response inhibition; the ability to purposefully reject reflex or dominant responses) compared to the control (Liu-Ambrose et al., 2008). There was also a significant falls rate reduction in the intervention arm and the authors concluded that improved executive function might be the method by which the OEP reduces falls, given that reduced cognition is a risk factor for falls. It may be that the OEP reduced falls via improved muscle strength and balance, and that improved cognition, whilst not being the mediator of reduced falls, is simply another positive outcome. This would need further investigation, possibly using strength and balance assessments as outcome measures.

Kyrdalen (2014) compared the use of OEP in groups versus performing it at home, in patients referred to a falls outpatient clinic. The intervention was 12 weeks in duration and the primary outcome, the Berg Balance Scale, was improved significantly in those performing the group exercise compared with the home-based training. There was also a statistically significant improvement in quality of life as measured by the SF-12 and in the 30-second chair rise test (Kyrdalen et al., 2014).

The OEP has proven to be consistently effective at reducing falls and fall-related injuries in community-dwelling older adults. Recorded adverse effects have been minimal, and the intervention has been shown to be cost-saving in the aged 80+ population. It is relatively cheap to administer and is simple and prescriptive in nature,

thus it is possible to ensure that patients receive an equitable programme across geographical sites, and indeed across countries. It is not surprising, therefore, that the OEP is the most commonly used falls prevention exercise intervention.

#### **1.10.2 Falls interventions: Falls Management Exercise**

Falls Management Exercise (FaME) was developed in the UK, originally funded by Research into Ageing. The authors developed what they called a '4-point plan', which described the overall goals of the exercise intervention; 1) to improve balance and coordination, 2) to increase functional capacity, 3) to strengthen bone and improve muscle mass and 4) to increase confidence thereby reducing fear of falling. They included seven types of exercise (dynamic balance, dynamic endurance, targeted strength training, flexibility training, getting down to and up from the floor (known as backward chaining), functional floor-work and tai chi) which had an existing evidence-base for falls prevention. The balance training evidence was from the Otago studies, and therefore FaME uses the same balance exercises as the OEP. As FaME also includes floor-based exercises, endurance exercises and tai chi, which are not part of the OEP, FaME is considered to be a more dynamic and challenging intervention than the OEP.

The original FaME study was published by Skelton and colleagues in 2005 and recruited community-dwelling, frequent falling (3 or more falls in the previous year) women aged 65 years and over for a 36-week group (class) exercise intervention (FaME) or a seated home exercise programme (control) (Skelton et al., 2005). The attention control exercises consisted of seated stretching and joint loosening activity that was deemed unlikely to impact on postural stability and therefore falls rate. Subjects were required to complete prospective falls diaries (returned fortnightly) for at least 36 weeks prior to the intervention start as well as throughout the intervention period and for at least 36 weeks post-intervention. The pre-intervention falls data were collected to allow the calculation of mean falls rate at baseline and this was 0.09 in both the exercise and control groups showing effective randomisation of participants by falls rate.

The authors reported a 31% reduction in the number of falls in the intervention group compared to the control (IRR 0.69, 95% CI 0.50-0.96). This was calculated using falls report data for the entire trial period but when the falls reduction was calculated in specific trial periods (during the intervention, following the intervention) they found that the reduction was entirely limited to the follow up period and its associated reduction rate of 54% (IRR 0.46, 95% CI 0.34-0.63). A further analysis of the follow up period for only those who had completed the intervention revealed a 34% reduction (IRR 0.66, 95% CI 0.49-0.90).

Although FaME has only been investigated in one trial, the reduction in falls was greater, over a shorter intervention period, than in the OEP studies. Plus, as FaME incorporates the OEP balance exercises, it could be argued that additional evidence for the potential effectiveness of FaME comes from the OEP trials.

### **1.11 Why study therapist effects?**

Therapist effects are potentially present whenever a treatment or intervention is mediated by a practitioner or therapist. If particular therapist attributes, characteristics or skills can be identified as having a positive influence on patient outcomes, therapist effects can be capitalised on to maximise intervention or treatment effectiveness via the selection, training and development of therapists. Therapist effects have been well studied in psychotherapy, less so in physiotherapy, and to date, only in one study of falls prevention exercise (see Chapter 2). Given that falls are common and costly to the NHS, reducing them is high on the health agenda. Exercise interventions to reduce falls need to achieve a dose of at least 50 hours (section 1.10). It may be that this can be reduced if therapist effects are utilised to good advantage.

#### **1.11.1 Definition and purpose**

Therapist effects have been well discussed, and indeed measured, in psychotherapy trials where it is acknowledged that therapists may have either a positive or negative effect on the intervention outcome. Therapist effects can account for variation in

standardisation of therapies. To assess the overall effectiveness of an intervention (treatment effects) when several therapists have been involved, it is arguably necessary to consider the therapist effect as a 'nuisance' factor for which the analysis needs to be adjusted. On the other hand, it may be that the characteristics of the therapist are of interest, and are therefore part of the research question. The method by which therapist effects in psychotherapy are dealt with from a statistical perspective has been debated (Wampold et al., 2010). Prior to the mid-1980s, therapist effects were largely ignored in psychotherapy studies, however, this can cause an overestimation of treatment effects (Wampold and Serlin, 2000). The Type I error hypothesis resulted in several studies that re-analysed data from previously published treatment effect studies in order to retrospectively investigate therapist effects (Luborsky et al., 1986, Blatt et al., 1996, Kim et al., 2006).

The size of therapist effects in psychotherapy has been more comprehensively studied in two meta-analyses (Crits-Christoph and Mintz, 1991, Baldwin and Imel, 2013). Therapist effect size is reported as a proportion of total variance in patient outcome that is attributable to therapists. Crits-Christoph and colleagues (1991) retrieved 15 studies and estimated mean therapist effects of 8.6%. Baldwin and Imel (2013) reported a mean therapist effect size of 5% from 46 included studies. In addition, both meta-analyses found that therapist effects were varied across studies and were greater in more natural settings versus within standardised interventions used as part of RCTs that frequently involved trial training of therapists and delivery of therapy according to manuals. The "homogenisation of therapists" (Wampold et al., 2010) within the strict trial setting is also applicable to the exercise research field, and is something that I will return to later in the thesis.

As well as the size of therapist effects, researchers have studied the characteristics of the therapist which may mediate therapist effects. Skill, expertise, competence and fidelity to the therapy model are all personal characteristics listed by Walwyn and colleagues (2010) as being potentially able to influence the content of, or perhaps adherence to, the therapy. More 'fixed' therapist factors (those that are not amenable

to change following additional training, for example), such as therapist age or gender, might also be of interest.

Okiishi and colleagues' large study involving over 5,000 patients who received counselling from 71 therapists investigated mostly 'fixed' therapist characteristics; therapists' years of experience, gender, type of training and theoretical orientation on patient recovery. They found no significant association between any of these characteristics and improved outcome for the patient (Okiishi et al., 2006). However, there was considerable variation in the number of treatment sessions, the rate of patient recovery and the overall amount of patient recovery suggesting that the therapists did potentially differ in their treatment effectiveness. This suggests that there may have been other therapist characteristics contributing to therapist effectiveness (ability to achieve the desired patient outcome) or that it was the patient characteristics (or the interaction between therapist and patient characteristics) that were responsible for the patient outcome.

In line with the concept that standard demographic characteristics, such as age and gender, are not predictive indicators of therapist effects, Anderson investigated interpersonal skills as well as the traditional demographic therapist variables, and found that older age (of the therapist) positively affected patient outcomes (Anderson et al., 2009). In a further analysis, the authors showed that the therapist age finding was explained by facilitative interpersonal skills (FIS) and that there was a significant association between FIS and patient outcome. In their rationale for investigating interpersonal skills, the authors state that "there is a need for therapist effects research that includes the measurement of constructs that are grounded in findings from psychotherapy research" (Anderson et al., 2009). Empathy and the ability to achieve a collaborative relationship with the patient are given as examples of these (and are assessed using FIS tools).

In line with the statement above regarding the interpersonal skills of psychotherapists, one might hypothesise that it is the exercise instructors' interpersonal skills that hold the key to therapist effects in the exercise setting too. There has been a limited



amount of research into the role of exercise ‘therapists’ or influence of exercise professional characteristics on outcomes in the exercise field. Exercise leaders’ impact on participants’ adherence to exercise has been more widely investigated (Hawley-Hague et al., 2016, Dinan, 2001, Ecclestone and Jones, 2004), for example, Annesi (1999) reported that the personal trait of ‘controlling’ in exercise professionals was positively correlated with clients’ adherence to exercise. Campbell (1997) comments that considerations for rolling out the Otago Exercise Programme as a public health programme should include “the enthusiasm and commitment of the research physiotherapist may encourage greater compliance in the elderly people than is possible in the busy routine of a working practice”. This clearly acknowledges that the personal skills of the person delivering the (falls prevention) exercise programme contribute to its success. Therapist effects in the field of exercise science and physiotherapy will be considered in greater detail in Chapter 2.

### **1.12 Why study dose?**

‘Dose’, when referring to exercise, is the term used to describe the total exercise intervention volume in hours (Sherrington et al., 2008, Sherrington et al., 2011). Dose of exercise does not describe exercise intensity, nor does it usually describe individuals’ levels of adherence to the intervention.

A dose of exercise of at least 50 hours is cited by falls exercise meta-analyses as contributory to effective falls prevention (Sherrington et al., 2008, Sherrington et al., 2011). This was calculated from trials of falls prevention exercise in patients with low and high falls risk, but a greater proportion of these trials (around two thirds) were in high risk patients. In the general, older population at low risk of falls, a lower dose, perhaps of more intense exercise, may be sufficient to positively affect falls outcomes.

Despite the best practice recommendation of at least 50 hours, Sherrington (2011) stated that there was not a clear cut-off point for dose in falls prevention exercise trials. I looked for trials with interventions of less than 50 hours in lower risk (general older adult) participants that were included in either the 2008 or the 2011 Sherrington

meta-analysis and found four that reported a reduction in falls (or falls risk factors) in the intervention group compared with the control. Voukelatos and colleagues reported fewer falls in relatively healthy community-dwelling older people attending 16 hours of tai chi compared with the control (Voukelatos et al., 2007). Carter and colleagues reported modest improvements in risks factors for falls (strength and balance) in young old, osteoporotic women after 40 hours of exercise classes compared with the control (Carter et al., 2002). Means reported a reduction in falls and fallers in ambulatory, community- dwelling older people allocated to 27 hours of exercise classes compared with the control (Means et al., 2005). Grahn Kronhed and colleagues (2009) showed that mean time to first fall in community-dwelling, osteoporotic women allocated to 30 hours of supervised exercise occurred later than in the control group (Grahn Kronhed et al., 2009). This suggests that in some low risk populations, falls can be reduced with between 16 and 40 hours of exercise.

An “augment” to the 2012 Cochrane Review of interventions for preventing falls (see section 1.8) looked at older peoples’ uptake of and adherence to the interventions (within trials) (Nyman and Victor, 2012). The median adherence to group-based exercise at 12 months post commencement of the intervention was at least 70%, but only 52% for individually targeted (home) exercise. Adherence as a moderator of trial outcome was reported in only five of the exercise trials. One of these reported a non-significant reduction in falls in the group who attended equal to or more than 75% of the exercise classes versus those who did not achieve 75% of classes (Lord et al., 1995). Further evidence regarding exercise effectiveness and adherence is highlighted by the trial of OEP with visually impaired older people (see section 1.10.1) where the authors found that those patients achieving three or more sessions per week had a 77% lower rate of falls than those exercising less than once per week (Campbell et al., 2005).

From the perspective of clinical practice, exercise trials need to investigate what dose of the intervention is sufficient to reduce falls *despite* the fact that there will inevitably be a range of patient adherence levels i.e. the dose offered in clinical practice should not be based on 100% adherence as this rate of adherence is unrealistic. However, as indicated by the trials discussed in the previous paragraph (Lord et al., 1995, Campbell

et al., 2005), the range of adherence within exercise studies makes it possible to conduct sub-analyses into different doses achieved as a result of different levels of adherence. In both these trials the purpose of the sub-analysis was to test whether the apparent ineffectiveness of the intervention was attributable to poor adherence. However, in trials of falls prevention exercise that are effective, similar adherence level sub-analysis could potentially be used to see if there is a dose cut-off point for falls outcomes lower than the 'full' dose under investigation. The definition of 'dose' used in my study will be discussed later, in Chapter 3, section 3.4.5.5.

### **1.13 Research questions and rationale**

Falls are very common in later life, with a high degree of associated morbidity. Multiple risk factors for falls have been described, and there are simple tests to identify those most at risk. Falls definitions vary widely in the literature and falls reporting is mainly by self-report and is potentially problematic. There is robust evidence for the role of exercise as a stand-alone intervention in the prevention of falls. The OEP and FaME interventions are the most widely used falls prevention exercise interventions in the UK and are recommended by the Department of Health (DoH, 2009).

As previously stated, the original FaME study recruited women who had fallen at least 3 times in the year preceding the study. This is a specific sub-group of the general older population and the results demonstrated that in this selected population, the rate of falls can be reduced by almost two thirds. We do not know if FaME is successful in reducing falls and falls risk factors in the general older population; community dwelling older people aged 65 or over, and in particular those at lower risk of falling. From a clinical significance perspective, should falls be reduced, it would support the use of the FaME intervention in the lower falls risk, general older population for the primary prevention of falls, in addition to the current approach; therapeutic exercise following falls or to treat high risk patients. We also do not know if a shorter intervention would be similarly effective as the original FaME intervention over nine months. This has clinical significance as it may reduce the cost of providing the

intervention, as well as potentially reducing wait lists in falls exercise services. We also do not know to what extent therapist effects may impact on falls rate and falls risk factors, nor if therapist effects can be capitalised on to maximise falls prevention exercise intervention effectiveness via the selection, training and development of therapists.

My work aimed to investigate the effect of the therapist and the effect of dose on falls outcomes within the Falls Management Exercise (FaME) programme. The aims, objectives and hypotheses for my study are described next, in section 1.13.1. As stated in the introduction to this chapter, my study uses data from a physical activity promotion trial, ProAct65+. A detailed description of the sections of the ProAct65+ trial that are relevant to my thesis can be found in Chapter 3, Methods. The primary falls outcome investigated in my work is falls rate. In order to learn how to analyse falls data (to help with my thesis) I led on the analysis of ProAct65+ falls data which was a secondary outcome in the trial. I describe the ProAct65+ falls findings in Chapter 4, Results 1, as they 1) demonstrate the falls analysis learning and 2) are foundation information for my study, bearing in mind that I am investigating the effects of exercise therapists and exercise dose on falls rate.

### **1.13.1 Aims, objectives and hypotheses**

**Aim:** The aims of this research were to study (a) the effect of the therapist (delivering the group exercise programme), and (b) the effect of exercise dose, on falls incidence and falls risk factors within a 24-week group exercise programme (FaME).

#### **Study objectives:**

1. The primary objective was to establish any difference in the number of falls over (a) the 24-week intervention period and (b) the 12-month follow up period, for subjects participating in the FaME exercise intervention, according to a) their allocated therapist and b) their intervention dose.

2. The secondary objective was to establish any difference in falls risk factors (balance and lower limb power) for subjects participating in the FaME exercise intervention, over the 24-week intervention period, according to a) their allocated therapist and b) their intervention dose.

### ***Hypotheses***

The over-arching *null hypotheses* were that (a) 'The therapist will not have a significant effect on the number of falls nor on the falls risk factors (balance and lower limb power).' (b) 'The intervention dose will not have a significant effect on the number of falls nor on the falls risk factors (balance and lower limb power).'

A priori, the working hypotheses for *therapist effects* were:

1. The age of the therapist will not have a significant effect on falls-related patient outcomes.

Rationale: Existing therapist effect literature (psychotherapists) suggests that age has no effect on patient outcomes. Older therapists may, however, have more clinical experience, so this characteristic was separately investigated (see hypothesis 4).

2. The gender of the therapist will not have a significant effect on falls-related patient outcomes.

Rationale: Existing therapist effect literature (psychotherapists) suggests that gender has no effect on patient outcomes. In falls exercise, however, male therapists may be perceived by patients as being more capable of 'lifting' the patient off the floor in the event of a fall, therefore giving the patient more confidence to exercise at a more challenging level which may, in turn, lead to improved falls outcomes. I think this is unlikely, but worth testing.

3. The professional background of the therapist will not have a significant effect on falls-related patient outcomes.

Rationale: Exercise professionals and physiotherapists have different skills but both approaches have strengths and therefore I see no reason why either group would achieve improved patient falls outcomes. From quality assurance observations, the most effective PSIs at each trial site did not have the same professional background. Also, the OEP literature reported training nurses to deliver the intervention and the nurses were more effective at reducing falls than their physiotherapist colleagues, suggesting that, following training, even those from backgrounds unrelated to exercise can deliver exercise interventions effectively.

4. Participants allocated to therapists with a greater number of years of experience of delivering the FaME intervention, compared with those allocated to therapists with fewer years of experience, will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: Familiarity with the typical patient group, and the intervention, potentially allows the therapist to a) tailor more effectively and b) adhere more faithfully to the evidence base. However, the physiotherapy literature did not provide evidence of therapist effects mediated by therapists' years of experience.

5. Participants allocated to therapists who achieved their PSI qualification before the trial, compared with those allocated to therapists who achieved their PSI qualification as part of the trial, will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: Therapists who attained their PSI qualification before the trial may have a greater number of years of experience of delivering the FaME intervention (see hypothesis 4 rationale).

6. Participants allocated to higher 'quality' therapists, compared with those allocated to lower 'quality' therapists, will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: 'Quality' of the therapist is the only measure in this analysis that encompasses several therapist 'characteristics', including; fidelity to the intervention, communication skills, motivation, observation and correction of patients, and as such, is more complex than the other factors analysed. As well as this 'quality' is likely to be reflective of some other single factors (such as previous clinical experience).

7. Participants allocated to therapists with high attendance, compared with those allocated to therapists with lower attendance, will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: Consistent attendance of the 'known' therapist is likely to positively affect patient adherence and therefore exercise dose. Existing literature suggests that therapists' characteristics can influence adherence to exercise.

8. Participants who enjoyed the exercises sessions, compared with those who did not enjoy them, will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: Enjoying the sessions is likely to positively affect adherence and therefore exercise dose. Dose attainment is already cited by falls exercise meta-analyses as contributory to effective falls prevention.

9. Participants who reported that the intensity of the exercises sessions was appropriate, compared with those who stated that the intensity of the exercises sessions was not appropriate, will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: Finding the sessions were at an appropriate intensity (not too easy, nor too hard) is likely to positively affect adherence and therefore exercise dose. Dose attainment is already cited by falls exercise meta-analyses as contributory to effective falls prevention.

A priori, the working hypothesis for *dose* was:

1. Participants achieving a higher intervention dose, compared with those achieving a lower dose will be 1) less likely to fall, and 2) more likely to achieve improved balance and lower limb power.

Rationale: A dose of exercise of at least 50 hours is reported to be necessary for effective falls prevention. This dose was calculated from trials of falls prevention exercise in patients with low and high falls risk. In the general, older population at low risk of falls, a lower dose may be sufficient to positively affect falls outcomes.

#### **1.14 Chapter conclusions**

This chapter has described the context for the research in this thesis; focusing on the problem of falls in the older population and the evidence for exercise, and in particular the evidence for the Otago Home Exercise Programme (OEP) and the Falls Management Exercise (FaME) programme, in the prevention of falls. I have introduced the aims of my study and the subject of therapist effects. I will expand on the latter in Chapter 2, in which I review the literature on the effect of therapists on patient outcomes in the context of exercise.



## **Chapter 2**

### **Literature Review**

#### **2.1 Chapter summary**

This chapter examines the existing literature relating to the effect of 'therapist' characteristics (such as rapport, communication, education, experience, age, gender) on patient outcomes following exercise or physiotherapy interventions. Psychotherapy studies show therapist effects (the proportion of total variance in patient outcome that is attributable to therapists) of 5–9% (Crits-Christoph and Mintz, 1991, Baldwin and Imel, 2013, Kim et al., 2006). Chapter 1, section 1.11 includes more detail regarding therapist effects in psychotherapy. The relevant literature relating to exercise 'dose' has been briefly described in Chapter 1, section 1.12.

In this chapter I will describe the methods employed for the systematic review, describe the studies retrieved and present a narrative synthesis of their findings. As scoping searches identified few relevant papers, the searches utilised were broad and not restricted to study type, population, setting or outcomes.

#### **2.2 Rationale for literature review**

As identified in Chapter 1, section 1.11, therapist effects have been identified and discussed in psychotherapy research and practice and it is plausible, therefore, that any intervention having a therapist at the interface between intervention and patient, will also benefit (or otherwise) from therapist effects. My hypothesis, therefore, was that exercise instructor or physiotherapist characteristics (for example, demographics, level of experience, intervention delivery skills) may influence patient outcomes during and following a dose (or prescription) of supervised exercise. This review attempted to test this hypothesis. The findings of the review were used to form the basis of the rationale for the main questions of my thesis, and to assist the interpretation of any findings from my therapist effect study, which focuses on possible therapist effects in a

falls prevention exercise intervention (FaME) in community dwelling adults aged 65 or over.

An interpretive approach to data synthesis was used as it was anticipated that both qualitative and quantitative studies of varied methodological quality would be identified. An integrative approach (such as meta-analysis) would only have been appropriate if studies were more homogeneous and comparable, and moreover met quality inclusion criteria, therefore allowing data to be aggregated and robust conclusions drawn. A narrative approach to synthesis of evidence (Dixon-Woods et al., 2005) was selected as this type of synthesis is suitable for diverse evidence, including both quantitative and qualitative studies. A thematic synthesis approach (Dixon-Woods et al., 2005, Thomas and Harden, 2008), which also benefits from its suitability to deal with diverse evidence types, was rejected on the basis that there was insufficient robust evidence found by this literature review to identify any recurrent themes. The narrative approach has been criticised for its lack of transparency (Dixon-Woods et al., 2005), however, the UK ESRC Methods Programme's guidance on the Conduct of Narrative Synthesis in Systematic Reviews (Popay et al., 2006) was consulted in order to understand current best practice guidelines for this approach.

This guidance identifies four elements to a narrative synthesis that were found to improve its transparency and reproducibility; “1) Developing a theory of how the intervention works, why and for whom, 2) Developing a preliminary synthesis of findings of included studies, 3) Exploring relationships in the data, 4) Assessing the robustness of the synthesis”. When the authors compared narrative synthesis (adhering to the guidelines) with meta-analysis, they found that implications for future research were elucidated more effectively from the narrative synthesis. Considering the paucity of literature on therapist effect in exercise interventions, this seems a highly appropriate method. However, the authors imply that only intervention or implementation studies can be synthesised using this model, so it was difficult to apply to my review, which identified studies which could neither be classified as intervention nor implementation (see Tables 2.1 to 2.4 for the range of study types identified). For

the above reasons, and despite its shortcomings, a standard narrative approach was adopted, as the only feasible approach for this review.

## **2.3 Methods**

### **2.3.1 Search Strategy**

A literature search of the following electronic databases was conducted; CINAHL (from 1982), EMBASE (from 1989), MEDLINE (from 1966), PsychINFO (from 1967), AMED (from 1985), to August 2015. The keywords used in combinations were: “therapist effect,” “therapist variation,” “therapist experience,” “therapist certification,” “therapist education,” “therapist training,” “therapist characteristics,” “exercise,” “physiotherapy,” “orthop(a)edics,” “musculoskeletal,” and “physical therapy.” These search terms were double-checked for effectiveness by checking that potentially eligible studies cited in papers already identified by the database search were independently included in that search. All potentially eligible studies were identified by reading the title and abstract to ascertain if it met the inclusion criteria. If this was unclear from scanning the title and abstract, the paper was included for further investigation. All papers considered to meet the inclusion criteria (or where there was some doubt) were retrieved for full-text screening. Those papers remaining eligible after the full-text screen underwent citation-tracking searches. Backwards, forwards, and lateral citation-tracking was used to search for any additional eligible papers (Greenhalgh and Peacock, 2005). Backwards citation-tracking was used to identify potentially eligible papers that were published prior to the paper being searched, by scanning the reference list. Forwards citation-tracking was used to identify potentially eligible papers that were published after the paper being searched, using the database’s electronic link to papers that cite the paper being searched. Lateral citation-tracking was achieved by using the database’s ‘related articles’ link. All papers

selected following the full-text screen were included in the review and underwent data extraction.

Inclusion criteria were:

- 1) the study was concerned with exercise, physiotherapy or physical therapy
- 2) the study examined the effect of at least one of the following therapist effects:
  - experience
  - education/training
  - personality/rapport/communication
  - patients' evaluation of therapist
- 3) the paper was published in English (for pragmatic reasons relating to lack of funding for translation)
- 4) the paper reported results from an original study (was not a review).

A study was excluded if:

- 1) the study examined therapist effects in non-exercise interventions e.g. massage, manipulation
- 2) the study examined therapist (trainer) effects in the education of therapists.

The search was not limited to RCTs. Dixon-Woods (2005) suggests that excluding qualitative studies from evidence synthesis, although historically common in systematic reviews, makes the evidence rather 'one dimensional' and might well omit important considerations such as feasibility and implementation. In line with this, Dixon-Woods suggests that some research questions can only be suitably answered by the inclusion of diverse forms of evidence. Therapist effects in exercise can potentially only be effectively examined in this way, not least, because there is currently very little published evidence, but also because only analysing quantitative data would not allow for the exploration of less tangible factors such as human personality, rapport, or other aspects of the relationship between patient and therapist which might affect patients' attitude or compliance with the intervention.

### 2.3.2 Data Extraction

The following data were extracted from the identified papers: general study information (title, lead author, country of study, year of publication), study characteristics (study type, population, type of intervention, type of therapist, therapist effect(s) studied, patient outcomes), and findings (overall interpretation of the results). I applied quality criteria but did not exclude studies on the basis of their quality assessment, for the following reasons:

Assessment of quality presents a challenge when dealing with diverse evidence. It has been suggested that different quality assessment tools may be utilised in one review, according to study type. Additionally, it has been suggested that quality assessment of qualitative papers is less necessary/desirable than that of quantitative studies (Dixon-Woods et al., 2005, Katrak et al., 2004), and indeed, some review authors have chosen not to assess qualitative study quality at all for fear of excluding important studies because of minor methodological flaws (Henderson and Ainsworth, 2000). Although there is still no consensus as to whether the quality of qualitative studies should be scrutinised in the manner conventional with experimental studies, the quality of qualitative studies can be assessed by a number of recognised tools. The Centre for Reviews and Dissemination's guidance for undertaking reviews in health care (2009) identified six critical appraisal tools designed for qualitative research. Several of these tools comprise a set of prompts for evaluation, whereas two are structured checklists; Quality Framework (Spencer et al., 2003) and Critical Appraisal Skills Programme (1998). Considering the paucity of literature on therapist effects in the exercise/physiotherapy field, and heterogeneity in the types of study included in my review, I decided not to use quality assessment to exclude publications. However, to assist with the interpretation of study results, quality criteria were applied to all studies included in my review.

I decided to assess the quality of papers using tools from one source. It could be argued that this results in the use of critical appraisal tools for specific study types that are not perhaps the most commonly used, for example; in a systematic review of RCT quality assessment tools, Olivo and colleagues (Olivo et al., 2008) reported that the

Jadad scale (Jadad et al., 1996) is the most frequently used. However, considering that studies were not excluded from my review on the basis of quality, and that quality criteria were applied only to assist with the interpretation of study results in narrative review approach, tools from one source were considered appropriate. Critical Appraisal Skills Programme (CASP, 2013) and National Institute for Health and Care Excellence (NICE, 2012), both provide checklists for a range of study types. I selected the NICE methodology checklists (NICE, 2012) because they are practical, feasible and 'user-friendly'. With regard to the quality assessment of RCTs, blinding of subjects to exercise is difficult (without a placebo exercise intervention), therefore, assessing risk of bias regarding blinding was focused on blinding of the outcome assessment and blinding of outcome assessors.

A quality assessment score was assigned to each publication. A maximum score of four was possible, with one point for each column of the NICE methodology checklist tables shown in section 2.4.7 (Tables 2.7 to 2.10). Studies scoring one point were considered to be of low quality, those scoring two or three were considered to be of moderate quality, and a score of four indicated a study that was considered to be of high quality.

### **2.3.3 Data analysis and synthesis**

The approach to data analysis and synthesis was a narrative review. It involves the ordering and reporting of findings in a descriptive manner to produce an account of the evidence. Narrative review can also involve analytic and interpretive reflection (Dixon-Woods et al., 2005).

## **2.4 Results**

### **2.4.1 Study selection**

133 papers were identified from the electronic database searches, 17 duplicates were removed, leaving 116 for screening. Of those, 106 were ineligible (63 were not exercise/physiotherapy/physical therapy interventions, 38 were exercise/physiotherapy/physical therapy interventions but did not analyse/discuss therapist effects, 4 concerned trainer effects in physiotherapy/physical therapy

education, 1 was a narrative review) and 10 were included (see Figure 2.1). Although the one identified review was ineligible, I proceeded with full-text inspection to cross-check for references, but no additional relevant references were identified. A further paper was identified from reading papers that cited the included publications, giving a total of 11 papers for full-text inspection. One of the 11 publications (Lewis et al., 2010) was found to have used data from three previously published RCTs so these papers were also retrieved for full-text inspection. However, they were excluded as they did not include therapist effect analysis/discussion. All of the remaining 11 papers were eligible for inclusion in the review. A PRISMA (Moher et al., 2009) flow diagram (Figure 2.1) summarises the search results and exclusions. Three of the retrieved publications were from the same research team (Resnik and colleagues) and reported data from the same population of patients and therapists, however, each paper reported a different analysis, therefore, I have discussed each publication separately in this review and referred to each of these three sub-studies as a 'study' in Tables 2.1 to 2.4 and 2.7 to 2.10. The 11 papers include studies that differ in design and include one randomised trial, one retrospective analysis from previously published RCT data, two qualitative studies, five cohort studies and two case-control studies (Tables 2.1 to 2.4). Study quality was assessed as described earlier and results of the quality assessment are reported later in sections 2.4.7 and 2.4.8.

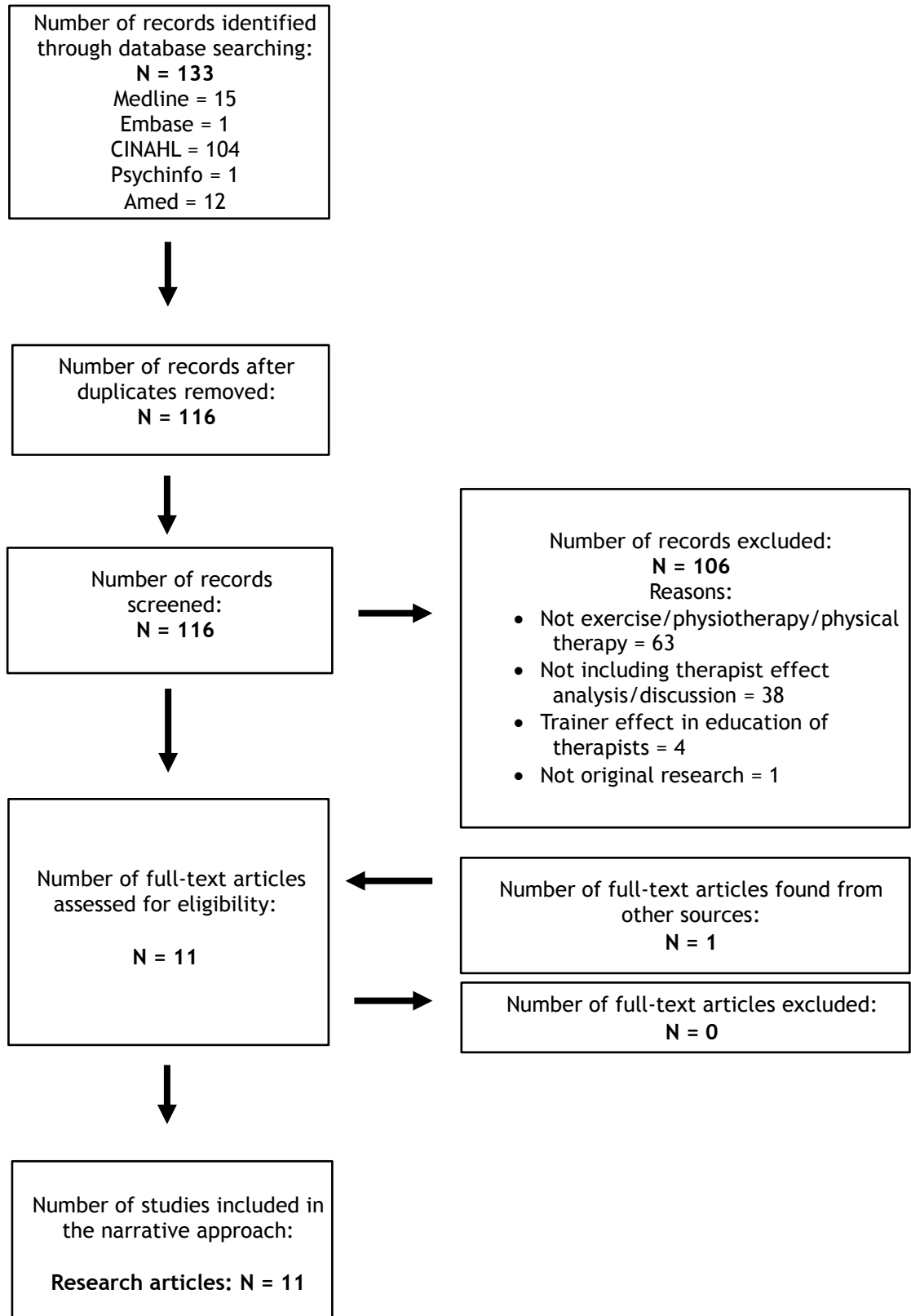
#### **2.4.2 Study Patient Population**

The number of subjects (patients) per study ranged from 56 to 24,276, although most had between 200 and 300 (Tables 2.1 to 2.4). The mean age of patients by study ranged from 34 to 76, although their age would most likely have been related to other inclusion criteria. For example, Hawley-Hague (2014) only included older adults (aged 60+) and several studies only recruited those with back pain, which is most commonly associated with middle-age. There was a higher proportion of women (91%) in the group exercise intervention study (Hawley-Hague et al., 2014) compared to the one-to-one interventions, which may reflect stereotypical exercise preferences with regard to gender and age (older women notionally prefer group exercise with an opportunity for

socialisation), otherwise the studies had attracted equal numbers of men and women. Tables 2.1 to 2.4 include a more detailed description of each study population.



**Figure 2.1**, PRISMA flow diagram describing the search process of finding articles exploring therapist effect in exercise/physiotherapy studies.



**Table 2.1,** Study description; RCTs

Title	Lead author & publication date	Country	Study Type	Population	Type of intervention	Type of therapist	Therapist effect studied	Outcome(s)
The Influence of Experience and Specialty Certifications on Clinical Outcomes for Patients with Low Back Pain Treated Within a Standardized Physical Therapy Management Program	Whitman JM, 2004	USA	Prospective randomised trial (secondary analysis)	n = 131 patients with low back pain. Mean age 33.9. 42% women	Physical Therapy	Physical therapist	Experience and specialty certification	Disability
Measuring practitioner/therapist effects in randomised trials of low back pain and neck pain interventions in primary care settings	Lewis M, 2010	UK	Data from 3 previously published RCTs (Dziedzic 2005, Jellema 2005, Hay 2005)	Dziedzic, Jellema & Hay, respectively: n = 350, 314, 402 patients with low back or neck pain. Mean age 51, 43, 41. 63, 48, 52% women	Physiotherapy (2 trials) and psychosocial intervention (1 trial)	Physiotherapist (2 trials) and GP (1 trial)	No specific mediator investigated	Self-reported disability and (Northwick Park Neck Pain Questionnaire (1 trial) and Roland and Morris disability questionnaire (2 trials), psychological outcome; SF12

**Table 2.2,** Study description; qualitative studies

Title	Lead author & publication date	Country	Study Type	Population	Type of intervention	Type of therapist	Therapist effect studied	Outcome(s)
Novice and Experienced Physical Therapist Clinicians: A comparison of How Reflection Is Used to Inform the Clinical Decision-Making Process	Flannery Wainwright S, 2010	USA	Qualitative	n = not reported. Patients receiving physical therapy following CVA	Physical therapy	Physical therapist	Novice/experienced therapist	Use of reflection during therapist-patient interaction
Using Clinical Outcomes to Explore the Theory of Expert Practice in Physical Therapy	Resnik & Jensen, 2003	USA	Qualitative; grounded theory	n = 24,276 patients with lumbar spine syndromes. Mean age 47.8 years	Physical therapy	Physical therapist	Years of experience, educational degree, specialty certification, gender, and practice setting	Patient self-reported health-related quality of life

**Table 2.3,** Study description; cohort studies

Title	Lead author & publication date	Country	Study Type	Population	Type of intervention	Type of therapist	Therapist effect studied	Outcome(s)
Influence of Advanced Orthopaedic Certification on Clinical Outcomes of Patients with Low Back Pain	Resnik & Hart, 2004	USA	Retrospective observational cohort	n=24,276 lumbar spine syndrome patients. Mean age 47.8. 58% women	Physical therapy	Physical therapist	Advanced orthopaedic certification	Health status measures were calculated: overall health status (OHS), SF-12 Physical Component Summary scale (PCS), and SF-36 physical functioning scale (PF-10)
Assessing the influence of treating therapist and patient prognostic factors on recovery from axial pain	Simon CB, 2013	USA	Prospective observational cohort	n=258 patients with axial pain. Mean age 46.4. 62 % women	Orthopedic physical therapy	Physical therapist	Years of experience	Visual analogue scale pain intensity, functional outcomes index
Multiple Levels of Influence on Older Adults' Attendance and Adherence to Community Exercise Classes	Hawley-Hague H, 2014	UK	Longitudinal observational cohort	n = 193 adults aged 60+. Mean age 76.1. 91% women	Exercise class	Exercise instructor	Age, gender, ethnicity, experience, training, personality, attitudes, professional background	Attendance and adherence
Physiotherapy intervention practice patterns used in rehabilitation after distal radial fracture	Bruder AM, 2013	Australia	Prospective observational cohort	n=75 patients with distal radial fracture. 71% aged over 51. 71% women	Physiotherapy	Physiotherapist	Years of experience	Physiotherapist records of type of physiotherapy intervention used and time spent administering intervention(s)

<b>Title</b>	<b>Lead author &amp; publication date</b>	<b>Country</b>	<b>Study Type</b>	<b>Population</b>	<b>Type of intervention</b>	<b>Type of therapist</b>	<b>Therapist effect studied</b>	<b>Outcome(s)</b>
Effects of Physical Therapist Training on Outcomes of Patients with Chronic Low Back Pain or Chronic Shoulder Pain	Levsen MJ, 2001	USA	Prospective observational cohort	n = 56 out-patients with chronic low back or shoulder pain. Mean age 50. No gender data	Physical therapy	Physical therapist	Training (long term post- graduate (PG) course)	Disability pain or Shoulder Rating questionnaire

**Table 2.4,** Study description; case-control studies

Title	Lead author & publication date	Country	Study Type	Population	Type of intervention	Type of therapist	Therapist effect studied	Outcome(s)
Influence of Orthopaedic Clinical Specialist Certification on Clinical Outcomes	Hart & Dobrzykowski, 2000	USA	Retrospective case-control	n = 258 patients treated in acute orthopaedic outpatient centres. Mean age 42. 65% women	Orthopaedic physical therapy	Physical therapist	Orthopaedic clinical specialist certification (OCS)	Changes in health status and efficiency as measured by visits, duration of treatment episode, and net revenue
Using Clinical Outcomes to Identify Expert Physical Therapists	Resnik & Hart, 2003	USA	Retrospective case-control	n=24,276 lumbar spine syndrome patients. Mean age 47.8. 58% women	Physical therapy	Physical therapist	Years of experience, advanced certification	Health-related quality of life

### **2.4.3 Intervention/therapy**

The type of therapy studied included routine physiotherapy/physical therapy delivered as usual (i.e. the exact treatment type and duration was not specified by the trial and was at the therapists' discretion), usual group-based exercise (i.e. the exact exercise type(s), duration and intensity was not specified by the trial and was at the exercise instructors' discretion) and standardised physiotherapy/physical therapy (therapy type is further detailed in Table 2.5). The preponderance of routine physiotherapy/physical therapy is consistent with study design; the majority of studies were observational. The heterogeneity of types of therapy made it impossible to directly compare findings, however, it highlighted a theme for consideration; therapist effects in standardised interventions versus therapist effects in routine physiotherapy/physical therapy. Only 2 studies reported standardised interventions; one reported a significant, yet small, therapist effect and the other reported no effect. Of the 9 papers describing non-standardised, routine therapy, 6 reported a therapist effect and 3 reported no effect. I will return to this point in the discussion.

### **2.4.4 Therapists**

There were 1101 'therapists' recruited in total across the 11 identified papers. (As described earlier, the same cohort of therapists was reported in the three publications by the same research team; Resnik and Hart, 2003, Resnik and Hart, 2004, Resnik and Jensen, 2003.) Seven studies recruited fewer than 17 therapists each. The number of therapists recruited per study did not appear to be related to the study type. There was a mix of therapist types, with the most common being physical therapists. Eight papers reported therapist gender, with 60% being female. Only 5 papers reported therapist age, the broadest age range (29-75) being amongst the group exercise instructors (who delivered exercise to the 60+ population). In line with therapist type, most therapists (10 papers) delivered one-to-one programmes. Only one study that recruited exercise instructors looked at a group-based intervention. Experience (of delivery/clinical practice) was very varied and ranged from 3 months to 43 years. Several studies specifically recruited therapists considered to be 'novice' to compare with those considered to be 'experienced'. Therapist characteristics are summarised in

Table 2.6.



**Table 2.5, Therapy type**

Publication	Type of therapy	Any control? What was it?	Frequency/Duration of therapy	Was the intervention standardised?	Delivery site	Therapist effect reported?
Whitman JM, 2004	Physical Therapy; exercise versus manipulation	Manipulation. 2 manipulation sessions and 3 stabilisation exercise sessions over 4 weeks plus daily home exercises	5 stabilisation exercise sessions over 4 weeks plus daily home exercises, no time per session reported	Yes. Specified exercises with pre-set progressions and set repetition numbers. Therapists trained by trial to ensure standardised practice.	Not reported	No
Lewis M, 2010	Physiotherapy; Dziedzic - advice, exercise and manual therapy versus advice, exercise and shortwave diathermy Hay - physiotherapy including manual therapy	Dziedzic - advice and exercise alone Hay - psychological pain management and exercises	Dziedzic - 8x 20-minute sessions over 6 weeks Hay - 1x 40-minute initial session followed by up to 6x 20-minute sessions	Dziedzic - Yes. Therapists trained by trial in agreed treatment protocols. Dose also standardised. Hay - Yes. Therapists trained by trial in agreed treatment protocols.	Dziedzic - physical therapy departments Hay - not reported	Yes
Flannery Wainwright S, 2010	Physical therapy	NA; not a trial	Not reported	No	Inpatient acute rehabilitation centres	Yes
Resnik & Jensen, 2003	Physical therapy	NA; not a trial	Not reported	No	Outpatient rehabilitation facilities	No
Resnik & Hart, 2004	Physical therapy	NA; not a trial	mean = 9 (SD=2) visits mean = 32 (SD=9) hours	No	Outpatient rehabilitation facilities	Yes

Publication	Type of therapy	Any control? What was it?	Frequency/Duration of therapy	Was the intervention standardised?	Delivery site	Therapist effect reported?
Simon CB, 2013	Orthopaedic physical therapies including manual therapy, mobilisation and exercise	NA; not a trial	Not monitored or measured	No	Outpatient facility	No
Hawley-Hague H, 2014	Multi-component exercise class including aerobic, strength, balance and flexibility exercises	NA; not a trial	Once weekly. On-going but study collected data for 6 months	No	Community venues	Yes
Bruder AM, 2013	17 physiotherapy interventions including advice, exercise, massage, and passive mobilisation	NA; not a trial	Mean consultation time = 30 minutes	No	Physiotherapy departments	Yes
Levsen MJ, 2001	Physical therapy	NA; not a trial	Therapist discretion	No	Outpatient clinics	Yes
Hart & Dobrzykowski, 2000	Orthopaedic physical therapy including exercise, massage and mobilisation	NA; not a trial	Therapist discretion	No	Acute orthopaedic outpatient centres	Yes
Resnik & Hart, 2003	Physical therapy	NA; not a trial	32 (SD=11) days for 'experts', 31 (SD=8) days for 'average group'	No	Outpatient rehabilitation facilities	No

**Table 2.6, Therapist characteristics**

Publication	Therapist type	Number of therapists	Gender n (%)	Age	Supervision ratio (therapist to patient)	Months/years of experience
Whitman JM, 2004	Physical therapists	13	2 (15.4%) women	Mean age 32.8 years	1 to 1	Average = 6.0 years of experience
Lewis M, 2010	Physiotherapists (psychotherapists excluded)	44	Not reported	Not reported	1 to 1	Not reported
Flannery Wainwright S, 2010	Physical therapists	6	3 (50%) women	26-40	1 to 1	3 = 96 months 3 < 12 months
Resnik & Jensen, 2003	Physical therapists	12	9 (75%) women	28-59	1 to 1	2.5-39 years
Resnik & Hart, 2004	Physical therapists	930	560 (60%) women	Not reported	1 to 1	0-43 years
Simon CB, 2013	Physical therapists	5	4 (80%) women	Not reported	1 to 1	1-13 years
Hawley-Hague H, 2014	Exercise instructors	16	14 (88%) women	29-75	Groups	3 to 120 months
Bruder AM, 2013	Physiotherapists	14	7 (50%) women	median age 33.5, 23-40	1 to 1	0.8-11 years
Levsen MJ, 2001	Physical therapists	6	Not reported	Not reported	1 to 1	At least 60 months

<b>Publication</b>	<b>Therapist type</b>	<b>Number of therapists</b>	<b>Gender n (%)</b>	<b>Age</b>	<b>Supervision ratio (therapist to patient)</b>	<b>Months/years of experience</b>
Hart & Dobrzykowski, 2000	Physical therapists	67	Not reported	Not reported	1 to 1	mean = 12.9 years specialists, 8.2 non-specialists
Resnik & Hart, 2003	Physical therapists	930	560 (60%) women	Not reported	1 to 1	0-43 years

#### **2.4.5 Therapist effect analysed**

Seven papers investigated a single therapist characteristic and four papers considered multiple therapist characteristics. One study looked for an association between patient outcome and therapist but did not investigate any specific explanatory variable/mediator of therapist effect. The range of characteristics analysed was as follows; years of experience (7 papers), advanced/speciality certification (7 papers), education/training/professional background (3 papers), personality/rapport/communication (2 papers), age (1 paper), gender (2 papers), ethnicity (1 paper), professional background (1 paper), attitudes to exercise (1 paper) (see Tables 2.1 to 2.4). It is not surprising that the most commonly investigated traits are experience and speciality certification, because 1) there is no evidence from the psychotherapy 'therapist effect' literature to suggest that therapist demographics explain differences in patient outcomes and 2) experience and speciality certification are 'modifiable' in the workplace. That is, if they prove to have an effect on patient outcomes there is a clinical practice implication that could be actioned (perhaps by providing additional training for therapists) which therefore justifies asking the research question. Only one study using an experimental methodology investigated more complex characteristics such as attitudes and personality.

#### **2.4.6 Outcome Measures**

Four papers in my review reported a single outcome and seven considered more than one. Outcome measures were the most diverse aspect of the identified literature. Most used 'direct' patient outcomes; disability (6 papers), health-related quality of life (5 papers), pain (4 papers), compliance/attendance/adherence (2 papers), change in health (2 papers), functional outcomes (1 paper), self-efficacy (1 paper) and one paper analysed the therapist's use of reflection during treatment sessions (see Tables 2.1 to 2.4). No studies reported falls outcomes.

#### **2.4.7 Quality of Evidence**

Quality of evidence is presented in tables according to study type, under column headings from the NICE Methodology Checklists. RCTs are shown in Table 2.7, qualitative studies in Table 2.8, cohort studies in Table 2.9 and case-control studies in

2.10. Quality scores were calculated and the scores and corresponding quality ratings are shown in Table 2.11. Quality of evidence is discussed for each study retrieved in my review in section 2.4.8; results of the narrative review, so that study findings can be appraised relative to study quality. Considering that quality criteria were not applied to retrieved studies in order to exclude studies from the review, it is especially important for study results and quality to be discussed together. Overall, study quality was varied; a few studies were found to be of high methodological quality and a few were found to be of low quality, however, the majority were considered of moderate quality.

**Table 2.7, Methodological quality of RCTs**

Study	Selection bias	Performance bias	Attrition bias	Detection bias
Whitman 2004	Low risk	High risk	Unknown risk	Unknown risk
Lewis 2010 (data from Hay 2005, Jellema 2005 & Dziedzic 2005)	Hay - Low risk Jellema - High risk Dziedzic - Low risk	Hay - High risk Jellema - High risk Dziedzic - High risk	Hay - Unknown risk Jellema - High risk Dziedzic - Low risk	Hay - Low risk Jellema - High risk Dziedzic - Low risk

**Table 2.8, Methodological quality of qualitative studies**

Study	Population (therapists)	Methods	Analysis	Relevance to guideline population
Flannery Wainwright 2010	Well reported	Well reported	Well reported and credible	USA. Novice and experienced physical therapists working in inpatient acute neurologic rehab settings.
Resnik & Jensen 2003	Well reported	Well reported	Well reported and credible	USA. Expert and average physical therapists working in varied practice settings but mostly orthopaedic outpatient.

**Table 2.9,** Methodological quality of cohort studies

Study	Selection bias	Performance bias	Attrition bias	Detection bias
Resnik & Hart 2004	High risk	Low risk	High risk	Low risk
Simon 2013	High risk	High risk	Low risk	Low risk
Hawley-Hague 2014	Low risk	Low risk	Low risk	Low risk
Bruder 2012	Unclear	Low risk	Low risk	Low risk
Levsen 2001	Low risk	Low risk	High risk	Low risk

**Table 2.10,** Methodological quality of case-control studies

Study	Research question	Selection of participants	Assessment	Confounding
Hart & Dobrzykowski 2000	Well covered	Not addressed	Poorly addressed	Poorly addressed
Resnik & Hart 2003	Well covered	Well covered	Well covered	Adequately addressed



The two RCTs included were found to have a high risk of performance bias (Table 2.7), due to the difficulty of blinding participants in exercise intervention studies. However, as discussed earlier in the chapter, this is a methodological problem shared by most trials of exercise/physiotherapy. Focusing more on the assessment of other types of bias, the Whitman (2004) trial provided low quality evidence due to some unreported detail (see section 2.5.1). The Lewis (2010) study reports on data from three RCTs (Hay et al., 2005, Jellema et al., 2005, Dziedzic et al., 2005) so these were separately assessed in order to establish the methodological quality of the Lewis study. Two of the RCTs (Hay et al., 2005, Dziedzic et al., 2005) were found to be of moderate to high quality, but the third RCT (Jellema et al., 2005) was assessed as having a high risk of 1) selection, 2) attrition and 3) detection biases (see section 2.4.8.3).

Both qualitative studies (Table 2.8) were well reported and found to provide high quality qualitative evidence.

The group of observational studies included in my review mainly comprised cohort studies. Cohort studies can either be based on routinely collected data that has already been collected and is retrospectively analysed (sometimes called retrospective, historical, or non-concurrent prospective, cohort studies) or on bespoke cohorts that were set up for the purposes of the study (sometimes called prospective, concurrent, or longitudinal, cohort studies). Song (2010) describes retrospective cohort studies as “a cohort of subjects selected based on exposure status is chosen at the present time, and outcome data, which were measured in the past, are reconstructed for analysis” (Song and Chung, 2010). In a therapist effect study of this type, for example, subjects who previously underwent physiotherapy for low back pain (for whom routinely collected outcome data is available) might be studied and categorised into cohorts by their exposure status; treated by a therapist with/without advanced certification. The outcomes of interest might be overall health status, and a physical functioning scale. Song (2010) states that a disadvantage of this design of cohort study is the limited control over the data collection, because data were gathered in the past and therefore may, for example, not include some variables of interest. Other considerations when assessing the methodology of cohort studies include 1) the exposure groups should be

from the same source population, 2) efforts should be made to minimise loss to follow up (in particular in prospective cohort studies with a long follow up period) and 3) findings should report any systematic differences relating to the outcome or exposure between those who drop out and those who remain in the study.

The NICE methodology checklist for cohort studies was used to evaluate the quality of the evidence, and in line with fact that the cohort studies included in my review were mainly 'retrospective' studies, 'treatment groups' was interpreted as 'exposure (to therapist) groups'. A summary of assessments of the methodology of cohort studies is shown in Table 2.9. The quality of studies was mixed with regard to selection, performance and attrition biases, although all studies were considered at low risk of detection bias. Quality of individual studies within the largest group of studies (cohort) in my review will be discussed in more detail in section 2.4.8.

A summary of assessments of the methodology of case-control studies is shown in Table 2.10. Of the two studies of this type in my review, Hart and Dobrzykowski (2000) had some limitations in internal validity when assessed using the NICE Methodology checklist for case-control studies (see section 2.4.8.2). Resnik and Hart (2003), however, was found to have sound internal validity.

#### **2.4.8 Results of the narrative review**

Overall in my literature review I found little evidence exploring therapist effects in exercise or physiotherapy. The most frequently investigated therapist characteristics were 1) advanced/speciality certification and 2) years of experience. Seven studies investigated each of these therapist characteristics (Tables 2.1 to 2.4). The studies exploring these therapist characteristics have been grouped in the narrative review. Due to the heterogeneity of studies, particularly in terms of design and patient outcome, no other themes were identified, therefore other studies have been described separately. A summary of all study findings can be viewed in Table 2.11.

**Table 2.11,** Summary of findings

Lead author & publication date	Therapist effect studied	Therapist effect reported?	Summary of Findings	Quality score	Quality rating
Whitman JM, 2004	Experience and specialty certification	No	Therapist characteristics did not contribute to clinical outcomes	1	Low
Lewis M, 2010	No specific mediator investigated	Yes	Therapist effects were accountable for between 3% and 7% difference based on disability, therapist effects were larger in interventions with a psychosocial approach	Hay 2 Jellema 0 Dziedzic 3	Moderate (2 studies) Low (1 study)
Flannery Wainwright S, 2010	Novice/experienced therapist	Yes	Reflection was used more frequently by experienced therapists	4	High
Resnik & Jensen, 2003	Years of experience, educational degree, specialty certification, gender, and practice setting	No	Extensive clinical experience is not necessary to achieve superior patient outcomes	4	High
Resnik & Hart, 2004	Advanced orthopaedic certification	Yes	Findings suggest a positive influence of therapist manual therapy certification on patient outcomes	2	Moderate
Simon CB, 2013	Years of experience	No	No statistically significant differences were found between therapists for either patient outcome	2	Moderate
Hawley-Hague H, 2014	Age, gender, ethnicity, experience, training, personality, attitudes, professional background	Yes	Instructor variables (age, gender, experience, training and personality) were positively associated with adherence	4	High

Lead author & publication date	Therapist effect studied	Therapist effect reported?	Summary of Findings	Quality score	Quality rating
Bruder AM, 2013	Years of experience	Yes	Junior therapists more frequently provided advice and taught a home exercise programme to patients compared with senior therapists	3	Moderate
Levsen MJ, 2001	Training (long term post-graduate (PG) course)	Yes	Mean number of treatment sessions needed to achieve same reduction in pain was statistically fewer for those therapists with PG training	3	Moderate
Hart & Dobrzykowski, 2000	Orthopaedic clinical specialist certification (OCS)	Yes	Patients treated by therapists with OCS received fewer visits compared with patients treated by therapists without OCS. Patient outcome was the same	1	Low
Resnik & Hart, 2003	Years of experience, advanced certification	No	No effect of therapists' years of clinical experience on patient outcome. No clear relationship between advanced certification and patient outcome	3	Moderate

#### **2.4.8.1 Therapists' years of experience**

The Flannery Wainwright study (2010) compared novice and experienced therapists' use of reflection during the therapist-patient interaction and found that reflection was used more frequently by the experienced therapists. The authors did not, however, investigate whether increased use of reflection in treatment sessions improved patient outcomes. This study was found to provide high quality qualitative evidence. Strengths of the study were;

- 1) study aims were clearly stated, as was a clinically relevant purpose,
- 2) only 6 therapists were involved, however, they were from three clinical environments, which reduced context bias,
- 3) respondents were an equal mix of gender and experience,
- 4) direct quotes from participants were included and these data were appropriately referenced to identify the sources of the quotes,
- 5) findings were credible and themed sub-headings were used to make the reporting of findings clear.

The limitations of this study were that the rationale for the selection of data collection and analysis techniques were omitted, making the research methodology less defensible, and some of the raw data were not reported. For example, the more experienced therapists were reported to discuss (in semi-structured interviews) 'reflection on professional experience' 3 to 4 times more than novice therapists, but the number of times in each group was not included.

The Whitman study (2004) investigated the effect of therapist experience and speciality certification on disability outcomes in patients with low back pain following either exercise alone or exercise and manipulation interventions. They reported no association of therapist characteristics with patient outcomes. However, the authors went to great lengths to ensure standardisation of exercise and manipulation techniques and suggest that this may have limited the use of clinical reasoning, and therefore the impact of therapist characteristics. They suggest that greater experience and qualifications are not necessary for a therapist to learn and deliver a strictly evidence-based, prescriptive exercise (or manipulation) intervention in the research

setting. With regard to intervention standardisation, it is worth noting that with a smaller number of therapists, intervention standardisation may be easier to achieve, but would also reduce any therapist effect estimates. There were only 13 therapists in the Whitman study. This trial provides low quality evidence because 1) the blinding methodology was not adequately reported; it was not clear if investigators were 'blind' to exposure to the intervention which raises concern regarding detection bias, and 2) because of potential attrition bias (differences between the treatment groups with respect to loss of participants) as no data were presented on the characteristics of non-completers and those for whom there were missing outcome data. The quality of this study means that any positive effect reported should be interpreted cautiously. However, no therapist effect was found.

Simon et al (2013) investigated the influence of the therapists' years of experience on patients' recovery from neck or low back pain. They reported no statistically significant difference between therapists for patient outcomes. They suggested that future research should focus on 'intrinsic' therapist characteristics such as beliefs and interests, as well as type and dosage of treatment, to investigate therapist effects. This study reported no significant differences between patients allocated to therapists at baseline. The methodological quality was assessed as being moderate. Those with missing follow up data were excluded from the analysis. Patients who were excluded were significantly different to those who were included, but this was not analysed by the treating therapist. It may have been that some therapists had a higher proportion of patients who did not return for follow up treatment because their back/neck problem had already resolved following treatment, or because their condition was aggravated by the treatment. This methodology was assessed as leading to a high risk of allocation bias. There was also a high risk of performance bias due to treatment type and duration being at the discretion of the treating physical therapist. Treatment type and duration were therefore confounders to the therapist effect analysis. The authors argue that their retrospective cohort study of therapist effects in routine, non-standardised physical therapy, therefore using different types of physical therapy, was

generalisable to 'real world' practice, however, it did not allow for therapist effects to be analysed separately from treatment effects.

Bruder (2013) studied physiotherapy following distal radius fracture and investigated the influence of therapists' years of experience on the type of physiotherapy intervention delivered and the time spent administering the intervention(s). They found that junior therapists more frequently provided advice and a home exercise programme compared to senior therapists, who were more likely to use 'passive' interventions such as joint mobilisation or massage. They point out that there has been little evidence since the late 1990s to support the use of passive interventions following wrist fracture, thus implying that the senior therapists may have been out-of-date in their treatment strategy. The authors conclude that "active interventions, including exercise and advice, were the most frequently administered interventions for patients after a distal radial fracture irrespective of physiotherapist factors" thus understating the therapist effect. The methodological quality of this study was assessed as being moderate. There was a low risk of attrition and detection biases, however, the baseline characteristics of patients by therapist group (those treated by junior therapists versus those treated by senior therapists) were not reported, therefore, the selection bias is unclear. Moreover, results may have been biased due to the fact that not all consultations were first consultations; around a third of them were follow up consultations on the same patient. The nature of the consultation (initial or follow up) could have influenced the duration or content of the session, however, authors were aware of this potential performance bias and therefore included a sensitivity analysis which demonstrated that follow up treatment sessions were similar to first treatment sessions. They could have avoided this potential bias by only analysing initial consultation data, however, they reported that the study would then have been underpowered (based on the reported sample size calculation).

My review includes three papers by Resnik and colleagues, all of which used the same data (and therefore therapists). The earliest publication (Resnik and Hart, 2003), approaches the study of therapist effect in an unconventional manner as therapists were not categorised a priori as experienced/novice or specialist/non-specialist

according to their baseline characteristics. Instead, patient outcomes (at the end of the observation period) were deemed to be the indicator of a therapist being 'expert'. The 'expert' group were then compared to the 'average' group in terms of years of experience and possession of advanced certification at baseline. They reported no effect of therapists' years of clinical experience on patient outcomes and no clear relationship between advanced certification and patient outcome. The internal validity of this paper was found to be moderate. The method of selection of participants (those treated by a physical therapist for low back pain, with intake and discharge data, in 1999 and 2000) from the source population (those treated for low back pain in 1999 and 2000) was well covered and the same exclusions were used for both cases and controls to reduce bias in the results. Differences between participants and non-participants were thoroughly investigated, however, some differences were reported which suggested the sample may not be representative.

In December 2003 the same research team published a qualitative paper presenting 'expert' therapist attributes (Resnik and Jensen, 2003). These included a patient-centred approach, a multi-dimensional knowledge base, collaborative problem solving with patients, use of movement observation, a passion for clinical care, valued continued professional learning and believing patient education was central to practice. The study appeared thorough in terms of the rationale for the methodological techniques used, and the data were 'rich'; for example, responses to questions were compared across the 'expert' and 'average' therapist groups. This study was considered to provide high quality evidence, however, it may be difficult to control for all patient factors that may have 'confounded' the therapist effect on clinical outcome.

#### **2.4.8.2 Therapists' advanced/specialist qualification**

The third Resnik and colleagues paper (Resnik and Hart, 2004) investigated advanced orthopaedic certification on outcomes for patients with low back pain. They reported statistically significant differences in all patient-reported outcome measures between patients treated by therapists with, and those treated by therapists without, advanced certification. They state that the results should be interpreted cautiously due to the



small number of therapists possessing advanced certification within the sample of therapists studied (40 therapists from the sample of 930 (4%)). This is the only research team who have used more than one analytical approach to explore the complexities of therapist effects. The methodological quality of this sub-study was assessed as being low due to those with missing follow up outcome data being excluded. Approximately a third of patients were excluded for this reason and systematic differences were found between those who were excluded (due to missing data) and those who were included, causing potential selection bias, and therefore reducing the generalisability of the findings.

Hart and Dobrzykowski (2000) studied the influence of therapist specialist certification on changes in patients' health status following treatment. They also reported on therapist 'efficiency' which was measured by the number and duration of treatment sessions, as well as net revenue. Patients treated by therapists with the specialist certification received fewer sessions compared with patients treated by therapists without specialist certification. Patient outcome was the same in both groups, thus the authors concluded that the more specialist therapists were more efficient and the treatment provided by them was less costly. They point out that a limitation of their study was the low number of therapists ( $n=67$ ). Compared with the other studies in my review, only one study involved more than 67 therapists. This study had some limitations in internal validity and was considered of low methodological quality; although the cases and controls were taken from the same (and therefore potentially comparable) population, the 'participation' rate for each group (specialist certification versus no specialist certification) was not reported, nor was a comparison between the cases and the source population. Representation of the eligible population amongst cases is therefore unclear. As well as this, it appears that 7 therapists with advanced certification treated the cases, however, 60 therapists treated the controls. The difference in treatment duration between specialist certification and no specialist certification groups reported in this study may have been caused by a clustering of

shorter treatment durations in the group with fewer (n=7) therapists. A broader spread of treatment durations may have been seen within a larger group of therapists.

Levensen et al (2001) compared the mean number of treatment sessions needed to achieve an equivalent reduction in patients' pain delivered by therapists who had attended a long-term post-graduate (PG) course compared with therapists who had not. They found that the average number of sessions was statistically significantly less for those therapists with PG training. This study had detailed patient inclusion and exclusion criteria that were used to ensure patients were as similar as possible at baseline for all characteristics other than exposure, to minimise confounding of results. Authors stated that the application of the detailed inclusion and exclusion criteria resulted in a small number of eligible patients. Despite this, a significant difference in baseline disability pain score was reported across groups. Change in disability pain score was analysed, thus adjusting for differences in baseline score; however, higher scores at baseline would have changed more than lower scores at baseline through regression to the mean, so some selection bias still applied. Performance and detection biases were assessed as low risk. Associated strengths in study design included 'blinding' of investigators to participants' exposure to therapists and other clinical factors (for example, baseline pain score) by using trained technicians to administer functional assessments and trained receptionists to consent patients and collect other baseline data. Risk of attrition bias, however, was deemed high due to the exclusion of patients post-treatment if they had not completed treatment or if outcome data were missing. The size of the attrition bias cannot be estimated as the number of patients excluded was not reported. Overall, this paper was considered to provide moderate quality evidence.

#### **2.4.8.3 Other therapist effects**

Lewis and colleagues (2010) conducted a complex analysis of data from three previously published randomised trials of interventions for low back pain or neck pain (Dziedzic et al., 2005, Jellema et al., 2005, Hay et al., 2005) that all showed no statistically significant differences between the treatment arms for their primary outcomes. They had, however, all shown improvement in the main disability outcome

across the entire study population. Following an analysis to calculate the variance between therapists as a percentage of the total sample variance, Lewis and colleagues (2010) reported an association between differences in disability outcome and therapist. Therapist effects explained about 3-7% of patient outcome differences and the effects were largest in interventions with a psychosocial approach. This study reanalysed data from three RCTs; two of which were found to be of moderate methodological quality (Table 2.11), but Jellema 2005 was assessed as having a high risk of 1) selection, 2) attrition and 3) detection biases for the following reasons;

- 1) the treatment groups may not have been similar at the start of the study because all patients treated by the same GP were allocated the same intervention, plus the patients were already aware of the allocation prior to consent,
- 2) no comparisons were reported between those who completed/did not complete treatment, nor between those with/without missing outcome data,
- 3) the GPs providing the treatment were responsible for collecting outcome data which may lead to bias towards a positive effect i.e. more patients may have been recorded as having achieved the outcome.

However, as far as the therapist effect analysis was concerned, there was successful 'blinding' of all participants, researchers and therapists as the data were collected for the primary investigations into treatments for back or neck pain, not for the secondary (therapist effect) analysis. Lewis (2010) reported practitioner effects in two of the three studies, one of which was considered to be of moderate quality and one that was considered to provide low quality evidence, therefore, overall, the evidence for practitioner effects from this study is limited.

#### **2.4.8.4 Therapist effects in group exercise**

The only study investigating group exercise delivered by exercise instructors, (Hawley-Hague et al., 2014), found that different instructor characteristics were associated with adherence at different time-points. Male gender, a greater number of years of experience, being younger, having attended motivational training and being perceived

as “conscientious” were positively associated with participant attendance at 3 months. Having an “extravert”, “agreeable” or “intellectual” personality was negatively associated with participant attendance at 3 months. At 6 months, only two personality traits were associated with attendance; “extravert” personality remained negatively associated and “conscientious” personality remained positively associated with attendance at this time-point. The assessment of instructor personality used a published scale (Saucier, 1994) including five traits; extraversion, agreeableness, conscientiousness, emotional stability and intellect. This study additionally analysed the exercise class participants' characteristics including demographics, health, socio-economic factors and attitudes to exercise, and found associations between a range of these and adherence to the exercise sessions. The paper acknowledges the complexity of the interplay between subjects' characteristics and 'therapist' characteristics in determining adherence to exercise programmes.

This prospective cohort study was rated as being of high methodological quality. Risk of selection bias was low as all participants who consented were allocated to therapist groups and the groups all received once-weekly multi-component group exercise for 6 months, making performance bias unlikely. The exercise was not standardised in content, however, this potentially does not matter given that the primary outcome was adherence; adherence is not associated with a specific exercise prescription and is therefore less ‘sensitive’ to exercise type than other patient outcomes, such as falls reduction. Attrition bias was also assessed as being low, given that attrition (non-adherence) was the focus of the study, and therefore the characteristics of non-adherers were investigated separately to therapist characteristics in a multi-level modelling analysis. The therapists were responsible for recording attendance data, and records could have been inaccurate or missing, however, the authors calculated a percentage attendance of all sessions offered in the investigation period, so missing data, at least, could have been adjusted for. The only limitation of the study appears to be that participants did not all start the exercise at the beginning of the investigation period, therefore, at baseline some participants could already have been attending for more than six months; a period that, as stated by the authors, is crucial in determining

ongoing exercise adherence. There is therefore potential that this confounds detection of the primary outcome (adherence) and may make a type 1 error more likely.

## **2.5 Discussion of findings**

### **2.5.1 Summary of findings**

There are very few studies on therapist effects in exercise or physiotherapy and the existing literature is heterogeneous with very variable methodology, outcomes and intervention type, and is of variable quality. The most frequently investigated characteristics were years of experience and advanced certification. Only one study using an experimental methodology investigated more complex therapist characteristics such as attitudes and personality. This is a gap in the research. Four of seven papers that investigated advanced/speciality certification reported a therapist effect (Table 2.11). Five of the seven were considered to be of moderate or high methodological quality. Three of seven studies that investigated years of experience reported a therapist effect (Table 2.11). Six of the seven were considered to be of moderate or high methodological quality. It is noteworthy that three studies reporting therapist effects showed differences between groups of therapists, for intervention content and length, rather than patient outcomes. This will be discussed further in section 2.5.3. Overall, seven of the eleven papers identified demonstrated a positive therapist effect result. However, quality assessment indicated that amongst those papers that demonstrated a therapist effect there was variable methodological quality and only two studies were considered to provide high quality evidence. The four publications that found no therapist effect were also of variable quality and only one of these was considered to provide high quality evidence. Overall, the key conclusion from my literature review is that further research would be worthwhile given that more than half of retrieved papers, albeit of variable quality, demonstrated a positive therapist effect.

### **2.5.2 Findings in context of other therapist effects literature**

In Chapter 1, section 1.11 I discussed literature on therapist effects in psychological therapies. Meta-analyses suggest that therapist effects in this field are in the region of

5-9% (Crits-Christoph and Mintz, 1991, Baldwin and Imel, 2013). This means that differences in patient outcomes that are attributable to therapists are 5-9% of the total differences in patient outcomes, and suggests that effect sizes are relatively small. However, there was a wide variation in therapist effects across studies, with larger effect sizes being found in studies of routine practice compared with RCTs using standardised interventions that involved trial training of therapists and delivery of therapy according to manuals. This theme has been identified in my review of therapist effect studies in physiotherapy, although the evidence is weak due to the paucity of studies and some poor methodological quality. More studies are needed to corroborate the hypothesis that physiotherapist effects are more likely or larger in studies of routine practice.

Another concept from the psychotherapy literature was that standard demographic characteristics, such as age and gender, are less likely than characteristics that are amenable to change, such as advanced knowledge and experience, to be predictive indicators of therapist effects. I did not find any robust evidence from the physiotherapy/exercise literature to suggest that this is not also the case in physiotherapy/exercise too. However, this may be due to the fact that there is little clinical reason for studying therapist characteristics that cannot be changed (by staff training, for example), nor could be used to recruit therapy staff.

### **2.5.3 Meaning and implications**

Whitman (2004) used standardised therapy techniques and suggested that this may have limited the impact of the therapist. This is of interest as the exercise programme used in the trial in which my study is nested is prescriptive and we attempted to standardise the intervention using similar methods to Whitman and colleagues; trial training, on site observations of therapists and standardised progressions. However, it also reinforces the need for further research into therapist characteristics which would

still prevail despite standardised, evidence-based exercise delivery, such as personality, rapport, empathy, or approachability.

The possibility of more qualified or experienced therapists achieving the treatment outcome in a shorter time-frame is also interesting. The studies demonstrating this effect involve therapists delivering one-to-one interventions where discharge is at the discretion of the therapist and when agreed treatment goals have been achieved. When the intervention length is not tailored to the individual; for example, in standardised group exercise interventions, might more highly qualified or experienced therapists have a more potent effect on outcome? Or might their potentially more effective sessions mean that even those patients with poor attendance benefit? An exploration of the combined influence of intervention dose and therapist characteristics might help explain this.

The qualitative studies included in my review considered more subjective therapist qualities or characteristics that had not been investigated in other types of study; for example, therapist experiences and thought processes. This is in line with the opinion that some things cannot be measured quantitatively. In a future bespoke therapist effects study, it might, therefore, be interesting to use qualitative approaches to investigate therapist characteristics that would be difficult to measure quantitatively. This was not possible in my study as data collected were pre-determined by the main trial's outcomes, which did not include an exploration of therapist effects.

Although the Hawley-Hague (2014) study appeared to be the most relevant study to my therapist effect study, it is noteworthy that the 'patient outcome' or dependent variable is attendance. This is due to the Hawley-Hague study population not being patients with specific treatment goals, rather, exercise class attendees. Attendance could be described as a 'participation' outcome, rather than a 'treatment' outcome, such as reduced back pain. The dependent variable in my study is the number of falls, which one could argue is a treatment outcome from a specific and proven exercise prescription. Attending any type of exercise will not reduce falls and indeed some exercise increases falls (Ebrahim et al., 1997). What is still unknown is whether any

reduction in falls through attending the required dose of standardised, evidence-based falls prevention exercise, is mediated by therapist characteristics.

The existing literature has also posed a question; might professional background, and associated therapist skill sets, be factors accounting for therapist effects? Lewis (2010) suggests that training non-psychologists to deliver psychologically-oriented interventions may explain greater variation in therapist effect. In falls prevention group exercise delivery, the 'instructors' are usually either physiotherapists or specialist exercise professionals. Whilst their appropriateness for the job is not contested, their skill sets are often different; exercise professionals tend to have more experience of delivering group exercise, whereas physiotherapists excel in exercising clinical populations and adapting exercise for individuals' specific disease-related impairments (tailoring). The FaME intervention, as already highlighted, is a group-based programme for frailer, older people at risk of falling, who, attributable to their age, frailty and falls status, are likely to have several co-morbidities, thus those delivering FaME require both group management skills and tailoring skills. An analysis that groups FaME 'therapists' by their professional background (therapist or exercise instructor) might elucidate this.

Variation in patient outcome may also result from therapists failing to adhere strictly to the evidence base. This phenomenon is referred to as 'therapist drift' by Waller (2009) and can be defined as changes to, or deviations from, an established, evidence-based treatment protocol. This 'drift' was observed by the author (SG) in clinical practice at Merton Sutton and Wandsworth Health Authority's Falls and Injury Prevention Exercise Service and appeared to occur more frequently when exercise instructors worked in isolation in community venues and did not receive continuing professional development (CPD) training. As a result of lack of standardisation, the Falls and Injury Prevention Exercise Service adopted an annual appraisal system,



allowing each instructor to receive feedback and action points regarding their practice and relating directly to the evidence base and/or best practice recommendations.

In the trial in which my study was nested, four quality assurance visits to group exercise instructors were conducted by researchers across each 24-week intervention. This, along with trial exercise instructor training, was to try to achieve standardisation of the intervention and avoid 'therapist drift'. Lewis (2010) concurs that therapist effects may be more noticeable in interventions that are less prescriptive, however, it could be argued that any variability in patient outcome following a non-prescriptive intervention, is actually a result of variability in the intervention. To effectively assess therapist effects in exercise studies, the intervention itself should therefore be as standardised as is possible. This was the case in Whitman (2004) where a multi-levelled standardisation approach was adopted. Therapists received trial training in the intervention techniques (manipulation and stabilisation exercises) by the same investigator. The same investigator also observed each therapist performing the techniques. Additionally, each treatment centre was supplied with a manual including, amongst other things, instruction on exercise progression. As well as this, intervention 'logs' were used by therapists for each treatment session to prompt therapists to adhere to the intervention programme. As previously stated, the trial in which my study was nested, used quality assurance visits and trial exercise instructor training, as well as providing exercise instructors with standardised session plans to prompt them to adhere to the intervention. This has allowed me to study therapist effects within a standardised treatment, thus reducing treatment variation which may decrease my ability to detect therapist effects.

#### **2.5.4 Strengths and limitations of the review**

Strengths of my review include that, to my knowledge, this was the first systematic review of therapist effects in falls prevention; systematic searching of five databases was employed, standardised methods of data extraction were utilised and quality criteria and ratings were applied. Limitations include that for pragmatic reasons the review was limited to English studies, only one reviewer judged the studies against the

criteria, I was only able to perform a narrative review due to heterogeneity of studies and conclusions were limited by the paucity of literature.

## **2.6 Chapter conclusions**

The aim of this review was to examine the existing literature relating to the effect of 'therapist' characteristics (such as experience, certification, rapport, communication, education, age, gender) on patient outcomes following exercise or physiotherapy intervention.

Eleven papers meeting the selection criteria were identified. As would be expected given that the search was not limited to a particular type of study, there was variety in study type. The studies were also varied in terms of methodological quality, the therapist characteristic analysed and the type of intervention. Moreover, patient outcomes were very variable (and included reduced pain and adherence to the therapy/exercise programme). There was insufficient methodological similarity between studies to make valid comparisons in outcome, however, there was some moderate and high quality evidence to suggest that therapist characteristics may contribute to the success of exercise/physiotherapy/physical therapy interventions.

There is a paucity of literature on therapist effects in exercise or physiotherapy interventions, yet psychotherapy studies show effects of 5–9% (Crits-Christoph and Mintz, 1991, Baldwin and Imel, 2013, Kim et al., 2006). The few, and methodologically diverse, papers retrieved by this review suggest that it is a promising area of research, but that many more studies are required to explain the complex interactions between therapist and patient characteristics and their effect on patient outcomes in exercise programmes. There is currently only one published study of therapist effects in the exercise setting which included falls prevention exercise instructors in the cohort of 'therapists'. There are no existing therapist effects studies that have focused solely on falls prevention interventions nor any studies that have reported the effect of the exercise therapist on falls outcomes.

## **Chapter 3**

### **Methods**

#### **Statement of Intellectual Property (IP)**

My thesis used data collected as secondary outcomes of the ProAct65+ study. This trial is described later, in section 3.2. Since the patient outcome investigated in my therapist effect study was falls, it was essential for me to understand the ProAct65+ falls data and analysis, and be able to interpret the falls analysis results. For this reason, I led on the ProAct65+ falls analysis and was first author of the ProAct65+ falls paper (submitted for publication). The IP for the falls analysis, however, remains with the ProAct65+ team. My thesis research questions on therapist and dose effects were not part of the ProAct65+ protocol, were my original idea and were developed by me for this thesis.

#### **Ethical Approval**

Ethical approval was granted to the ProAct65+ trial from Nottingham Research Ethics Committee 2 (application number 08/H0408/72). National Health Service Research & Development approval was granted by NHS Nottingham City, Nottinghamshire County, Derby City, Derbyshire County and Westminster, Brent, Harrow, Hounslow and Barnet & Enfield Primary Care Trusts. The therapist and dose effects study presented in this thesis was nested within the ProAct65+ trial and ethical approval was therefore covered by the main trial's approvals. Written informed consent to participate in the study was obtained by all 'participants'. The 'participants' included all patients and therapists whose anonymised data were used in the therapist and dose effects study.

### **3.1 Chapter summary**

As described above, my study is embedded within another study; the ProAct65+ trial. This was a multi-centre cluster-randomised controlled trial that compared two exercise interventions (Falls Management Exercise (FaME) and the Otago Exercise Programme (OEP)) with usual care in community-dwelling over 65s recruited through general practice. The full Health Technology Assessment (HTA) report (Iliffe et al., 2014) can be

accessed from <http://www.ncbi.nlm.nih.gov/pubmed/25098959>. The primary objective was to investigate the effect of the interventions on longer term exercise habits. Falls data were collected and indeed the effect of the interventions on falls incidence was one of several secondary objectives. As described at the end of the chapter 1 (section 1.13) my study investigates the effect of the therapist and intervention dose on falls outcomes. This chapter therefore covers the methods relating to three key areas:

- ProAct65+ methods that relate to my study, including the recruitment of patients, the recruitment and training of the specialist exercise instructors and the protocol of the exercise interventions.
- ProAct65+ methods that relate specifically to falls data and the falls analysis, including the falls outcome measures and the analysis plan.
- The methods of my therapist and dose effects study, including the aims, data management and analysis plan.

## **3.2 ProAct65+**

### **3.2.1 Design, aims and objectives**

#### **Aims and objectives**

The ProAct65+ Study was a cluster-randomised controlled trial that compared two exercise programmes, FaME and OEP, with usual care (no specific exercise programme provided) in community-dwelling adults recruited through general practices in London and Nottingham (see Appendix 1, the ProAct65+ protocol paper (Iliffe et al., 2010) for further details). The exercise interventions were 24 weeks in duration and the participants were followed up for up to two years post-intervention. The primary outcome, minutes of moderate to vigorous intensity physical activity (MVPA) per week at one year post intervention, showed a significant improvement in the FaME group compared to the control, but not in the OEP group (see Appendix 2, the ProAct65+ primary outcome paper (Iliffe et al., 2015) and Chapter 4, section 4.6 for a summary of the findings). Although the primary outcome was not falls, the exercise regimes utilised were falls prevention programmes, and the general older population recruited was expected to contain older people who had already fallen and those who displayed

other risk factors for future falls, such as poor balance and poor leg strength. A reduction in falls rate was therefore a secondary outcome. Other secondary objectives included determining the health benefits of the two exercise programmes to participants, assessing the cost-effectiveness of the interventions compared with treatment as usual and assessing the acceptability of the programmes to participants.

## **Design**

The ProAct65+ study had three arms; a group-based exercise intervention (FaME), a home exercise intervention (OEP) and treatment as usual (the control). The study had a parallel design. The trial is described as cluster-randomised meaning that general practices were randomised to the three trial arms, and all participants recruited from a practice were therefore all allocated to that arm. Patients were recruited from primary care in one of two sites; London and Nottingham/Derby. Randomisation to study arm was by practice using minimisation. Minimisation aims to ensure treatment arms are balanced, not only for the number of patients in each group, but also for specific pre-set factors. The variables used in the ProAct65+ minimisation process for were trial site, practice size and practice deprivation score. Patients were followed up for up to 2 years post intervention to determine the impact of the exercise programmes on longer-term continuation of exercise.

### **3.2.2 FaME and OEP interventions**

The trial had 3 arms: the home-based exercise programme (OEP), the community venue-based group exercise programme (FaME) and usual care (UC). The exercise interventions mirrored their previous trials (see Chapter 1, sections 1.10.1 and 1.10.2), but with some modifications. These included a shorter duration (24 weeks) and volunteer Peer Mentors (PMs) were recruited to provide support (home visits and telephone calls) to OEP participants. Full compliance in the exercise programmes totalled 48 hours and 36 hours in the FaME and OEP groups, respectively.

#### **3.2.2.1 ProAct65+ Otago Home Exercise Programme**

The OEP is a home-based falls prevention exercise programme that aims to reduce falls by strengthening the lower limb muscles and improving balance. Outcomes of previous

OEP trials and associated intervention details can be viewed in Chapter 1, section 1.10.1. OEP subjects were inducted in the exercise technique in a group session (of up to fifteen participants) by a trained exercise professional (myself) and provided with a personalised exercise plan (suited to their baseline exercise tolerance) and a set of ankle cuff weights. Following this, they were encouraged to exercise at home unsupervised but received two home visits and approximately eight telephone calls from trained volunteer Peer Mentors (PMs) to aid compliance. Each bout of exercise was designed to last approximately 30 minutes and subjects were required to perform their set of exercises thrice-weekly, thus totalling 36 hours (of strength and balance retraining exercises) in the 24-week intervention period (assuming 100% compliance). An attempt was made to progress the intensity of the exercises over the intervention period by increasing the weight of the ankle cuffs and by reducing the external support used when performing the balance exercises. Ankle weights used by participants ranged from 1kg to 4kg. Those in the OEP group were also asked to take two 30 minute walks per week. The intensity of the OEP is described by the original authors (Campbell and colleagues) as moderate. Participants were asked to record the duration of exercise and the number days per week they carried out the programme using exercise diaries. Table 3.1 identifies the differences between the original OEP study intervention and the ProAct65+ OEP intervention. The key differences in programme design were the length of the intervention (1 year versus 24 weeks) and physiotherapist versus peer mentor support.

**Table 3.1,** Comparison between original OEP and ProAct65+ OEP

Original Studies (Campbell 1997, 1999a, 1999b, 2005, Robertson 2001a, 2001b)	ProAct65+
1 year	24 weeks
75/80+	65+
Physiotherapist supervision and physiotherapist/OEP trained nurse support	Exercise professional supervision and trained peer mentor support
6 home visits & 12 telephone calls	2 home visits and 8 telephone calls
Taught exercises in home	Taught exercises in group
Tailored programme	Tailored programme
Structured progression	Progression via Peer Mentor
Walking 2x 30mins per week	Walking 2x 30mins per week

### **3.2.2.2 ProAct65+ Falls Management Exercise (FaME) programme**

The FaME intervention was a once-weekly group-based programme led by a qualified postural stability exercise instructor (PSI) in a community venue. FaME sessions were planned to be attended by no more than ten participants, to allow for adequate supervision. This one hour supervised session was supplemented with twice-weekly bouts of a 30-minute home exercise programme, thus totalling 48 hours across the 24-week intervention. The sessions contained balance, muscular strength, cardiovascular endurance, flexibility and tai chi exercises, as well as an approach for getting down to and up from the floor (backward chaining) and functional floor-based activities. Exercises were progressive in intensity over the intervention period, and the intensity (especially of the balance exercises) at any given time-point was moderate to high. Those in the FaME group were also asked to take two 30 minute walks per week. Table 3.2 identifies the differences between the original FaME study intervention and the ProAct65+ FaME intervention. The key difference in programme design was the length of the intervention (9 months versus 24 weeks).

**Table 3.2,** Comparison between original FaME and ProAct65+ FaME

Original Study (Skelton 2005)	ProAct65+
9 months	24 weeks
65+ women	65+
Frequent fallers	Frequent fallers excluded
One-hour class per week	One-hour class per week
Taught by PSI	Taught by PSI
Up to 15 participants per class	Up to 10 participants per class
Home exercises x2 per week	Home exercises x2 per week
Tailored programme	Tailored programme
Structured progression	Structured progression – more advanced
No walking	Walking x2 30mins per week

### 3.2.2.3 Usual Care

Patients in the usual care (control) arm were not allocated an exercise programme, however, they were not discouraged from participating in their usual physical activities.

### 3.2.3 Sample size

The ProAct65+ trial was powered to detect effects of the interventions on moderately vigorous physical activity (MVPA). The sample size needed was 1200 patients from 30 general practices. The power calculation for ProAct65+ can be viewed in the full report (<http://www.ncbi.nlm.nih.gov/pubmed/25098959>). The trial was not powered to detect any effects on falls. This will be discussed further in section 3.4.3.

### 3.2.4 Recruitment of general practices

General practices were recruited through the National Institute for Health Research (NIHR) Primary Care Research Networks (PCRN) in the two trial sites. The PCRN were



requested to identify potential participant practices based on previous research participation and/or interest in exercise trials. Potential practices were sent invitation letters and telephone or personal contact was made with practice managers as necessary.

### **Practice inclusion criteria**

Practices were suitable for inclusion if (a) they were committed to participate over the entire study duration and (b) a community venue suitable for group exercise was available in the local area.

### **3.2.5 Patient selection, recruitment and consenting**

#### **3.2.5.1 Recruitment of participants**

Participants were aged 65 years or over, registered with participating general practices, living independently (not in residential or nursing homes) and physically able to take part in group exercise. Frequent fallers (3 or more falls in the past year) were excluded as were those already achieving sufficient exercise to benefit health. Other exclusions included uncontrolled medical conditions and significant cognitive impairment. Inclusion and exclusion criteria are listed below.

#### **Participant inclusion criteria:**

1. Aged 65 years or older
2. Registered with participating general practices
3. Gave informed consent to participate
4. Able to walk without help from another person (with a walking aid, if needed) and physically able to taking part in a group exercise class.

#### **Participant exclusion criteria:**

Patients were excluded by their GP practice or by the researcher at baseline meeting if they had any of the following:

- 1) Unstable/uncontrolled medical condition(s) (for a full list of medical exclusions see the HTA report)
- 2) Three or more self-reported falls in the year preceding the study (frequent fallers)
- 3) Psychiatric conditions or significant cognitive impairment which would prevent safe participation in an exercise class
- 4) Not able to maintain a seated upright posture or unable to mobilise independently
- 5) Not living independently (e.g. living in residential care)
- 6) Receiving physiotherapy at recruitment
- 7) Already achieving government recommended PA levels (5 x 30 minutes per week)

#### **3.2.5.2 Recruitment procedure**

A random sample of 600 eligible patients was invited to participate via a letter from their GP. In practices with less than 600 people aged 65 or over, all potentially eligible patients were invited to participate.

Patients who returned the expression of interest reply slip were called by a researcher to arrange a suitable time and date to meet the researcher at the GP practice. This meeting lasted about an hour and included gaining the patient's consent to participate, giving the patient trial information and gathering baseline data (sociodemographic, medical history, falls history, functional assessment scores and questionnaire data). A full list of outcome measures is listed later, in section 3.2.8. Following the completion of all baseline meetings, a list of consented patients was supplied to the GP practice so that the GP could advise the trial of any exclusions on the basis of medical ineligibility.

#### **3.2.6 Allocation concealment and randomisation**

Randomisation to study arm was by practice. The practices, their patients and the researchers were all blinded to allocation until all patients at a practice were recruited.

### **3.2.7 Blinding**

Blinding participants in exercise trials (beyond the recruitment meeting) is challenging. The use of an attention control is common in exercise intervention studies, for example, using a modality of exercise that is not thought to confer the patient outcome(s) under investigation. In falls exercise research, therefore, seated exercise that targets neither balance nor muscle strength might be appropriate. In ProAct65+ the use of an attention control would have been costly and impractical given the large size of the patient sample recruited, so for pragmatic reasons, usual care (no specified exercise programme) was appropriate. This pragmatic approach also meant that the trial was looking at the difference between groups in 'real life', therefore any social effect of attending group exercise, for example, did not need to be measured. However, the lack of an attention control meant that it was impossible to blind patients once the allocation of their GP practice to a trial arm was made. Researchers were blind to the allocation of a practice to a trial arm (and therefore also to participants' allocation) at baseline assessment, as were the practices. Disclosure of treatment arm allocation of a practice did not occur until recruitment at that practice was complete.

### **3.2.8 Outcome measures**

The details of all outcome measures, including their derivation and validation, can be seen in the full HTA report. The primary outcome concerned the volume of physical activity measured using the Community Healthy Activities Model Program For Seniors scale (CHAMPS) and the Physical Activity Scale for the Elderly (PASE) questionnaires, and the Phone-FITT. These were administered at baseline, then 0, 6, 12, 18 and 24 months after the end of the interventions, however, the primary analysis was of CHAMPS data collected at 12 months beyond the close of the interventions. The interventions were compared separately with the control group. I will describe the falls-related outcomes used in this thesis in more detail in section 3.3.2.

The outcome measures relating to the secondary outcomes were:

- Functional assessments (Timed-Up-and-Go (TUG), Functional reach (FR), Modified Clinical Romberg Static Balance test, 30 second chair rise) \*

- Confidence in Maintaining Balance (CONFbal) scale\*
  - Short Falls Efficacy Scale-International (Short FES-I) \*
  - Outcome Expectation for Exercise (OEE)
  - Older People's Quality of Life Questionnaire (OPQoL) & the Short Form questionnaire-12 items (SF-12).
  - Lubben Social Network scale (LSNS) & the Multidimensional Scale of Perceived Social Support (MSPSS)
  - Attitudes to Falls-Related Interventions Scale (AFRIS) questionnaire
  - Falls Risk Assessment Tool (FRAT)\*
- \*Falls-related outcome measures (see section 3.3.2)

### **Data collection time points**

Study data collection commenced in August 2009 and ended in September 2013. Face-to-face data collection was only scheduled at two trial time points; baseline and immediately following the 24-week intervention (0 months post-intervention), however, as there were a large number of outcomes and meetings with patients were planned to last no longer than an hour, some data were collected via self-completion booklets even at baseline and immediately post-intervention. All other data collection was via self-completed questionnaires and telephone interview at 6, 12, 18 and 24 months post-intervention. The functional assessments were assessed at baseline and immediately post-intervention only. Table 3.3 lists all instruments and collection points.

**Table 3.3,** Schedule of outcome measures

Measure	Method administered	Baseline	Immediate post-intervention (24 weeks from randomisation)	6 months post-intervention	12 months post-intervention	18 months post-intervention	24 months post-intervention
CHAMPS	Self-complete questionnaire	√	√	√	√	√	√
PASE	Self-complete questionnaire	√	√	√	√	√	√
Phone-FITT	Telephone questionnaire	√	√	√	√	√	√
Timed-Up and-Go	With researcher	√	√				
Functional reach	With researcher	√	√				
Romberg Static Balance test	With researcher	√	√				
30-second chair rise	With researcher	√	√				
CONFbal	Self-complete questionnaire	√	√	√	√	√	√
Short FES-I	Self-complete questionnaire	√	√	√	√	√	√
OEE	Self-complete questionnaire	√	√	√	√	√	√

Measure	Method administered	Baseline	Immediate post-intervention (24 weeks from randomisation)	6 months post-intervention	12 months post-intervention	18 months post-intervention	24 months post-intervention
OPQoL	Self-complete questionnaire	√	√	√	√	√	√
LSNS	Self-complete questionnaire	√	√	√	√	√	√
MSPSS	Self-complete questionnaire	√	√	√	√	√	√
AFRIS	Self-complete questionnaire	√	√	√	√	√	√
FRAT	Self-complete questionnaire	√	√	√	√	√	√

### **3.2.9 Recruitment and training of peer mentors and PSIs**

As a researcher in the ProAct65+ trial, and qualified PSI, I led on the training of the exercise instructors and peer mentors, in both OEP and FaME arms.

#### **3.2.9.1 Recruitment and training of PSIs**

Postural Stability Instructors (PSIs) are exercise professionals or allied health professionals (physiotherapists or occupational therapists) who are trained in the design and delivery of the FaME intervention at National Vocational Qualification (NVQ) Level 4; the highest level of exercise qualification, targeting clinical populations. The Postural Stability Instructor Award has been delivered by Later Life Training (LLT) (an independent, not-for-profit training organisation) for some 15 years, so there are already many trained PSIs working for NHS falls prevention exercise services nationally. ProAct65+ looked to recruit (where possible) PSIs who were already qualified and had clinical experience, but additionally had the resources to train PSIs, should it prove difficult to recruit experienced PSIs. It was planned to recruit and/or train 12 PSIs per trial site. PSIs were initially approached to deliver the shortened FaME intervention for ProAct65+ by email contact from a list supplied by LLT. Following this, exercise professionals working with active, independent, older adults, and physiotherapists, were approached to undertake the PSI qualification as part of the trial training and then work for the trial delivering the shortened FaME intervention. Both experienced and inexperienced PSIs also undertook one day's trial specific training to learn about trial data collection and the planned progression of the 24-week intervention. Trial PSIs were required to undergo a Criminal Record Bureau (CRB) check.

Session plans were supplied to the PSIs to ensure standardisation of the intervention across trial sites and between individual PSIs. PSIs were engaged to deliver a once-weekly session of an hour's duration (per group of allocated patients) and to facilitate socialisation and refreshments for around half an hour immediately after the session. PSIs were required to keep registers of attendance for all FaME sessions and to follow-up any patient absences from sessions by telephone call. The PSIs were supplied with exercise equipment, including resistance bands and mats, as well as trial Home Exercise booklets, which they were required to issue to patients within the first few

weeks of the intervention (and when the exercises contained within had been practiced within FaME sessions). PSIs were supplied with copies of the health assessment tool and functional assessment scores completed at the baseline appointment with the researcher, so that they could tailor exercises to patients' medical and functional status. PSIs were requested to liaise with the research team regarding any patient drop outs or long periods of absence and to keep a record of any adverse events occurring within sessions on designated trial paperwork. Registers were returned to the research team each month along with PSI invoices. PSIs were responsible for reminding participants to fill in their exercise diary, including on it each FaME class and bout of home exercise performed.

### **3.2.9.2 Recruitment and training of PMs**

Peer Mentors (PMs) were recruited to support and motivate OEP participants within ProAct65+. PMs have been previously used as motivators within the context of physical activity promotion, but ProAct65+ was the first trial using PMs to support falls prevention exercisers. It was calculated that 40 PMs, aged 65 or over, would be needed per site, if each one mentored 5 patients. PMs were recruited through a variety of methods including newspaper advertisements, voluntary organisations, from leisure services older adult exercise classes and from the group of patients found to be ineligible to participate in ProAct65+ due to already achieving 150 minutes or more of MVPA per week. PMs were trained by trial staff over two consecutive days, during which they learnt about the mentoring role and associated responsibilities, as well as the technique for the OEP exercises. Trial PMs were required to undergo a Criminal Record Bureau check.

It was initially planned that PMs made four home visits to each patient and telephoned twelve times across the 24-week intervention period. This schedule was reduced to two visits and eight telephone calls (after the first practice was allocated to the OEP arm) in an attempt to boost poor PM recruitment (Stevens et al., 2013). PMs were required to complete trial Patient Contact forms, detailing the date, duration and content of any visits and telephone calls. Two trial staff met with groups of PMs at four scheduled meetings across the 24-week intervention period so that PMs could share



their experiences of mentoring and researchers could answer any questions. The meetings were also used to deliver some additional training to PMs regarding the progression of the OEP exercises. It was hoped that the meetings would help to standardise PM activity.

### **3.2.10 Quality control in the intervention arms**

As a researcher in the ProAct65+ trial I led on the quality assurance of the exercise interventions.

The intention to provide quality assurance and standardise the exercise interventions was included in the original ProAct65+ trial protocol, however no detail of how this should be achieved was included. I led on the development of the methods and associated documentation. The detail regarding the QA of the OEP intervention is included in this section (3.2.10.2), however, the FaME QA procedures and output data are detailed in section 3.4.5.2. This is because I extended the FaME QA documentation to include information on therapist characteristics and participant perception of sessions, which was used to produce data exclusively for my therapist effects study.

#### **3.2.10.1 FaME quality control summary**

See section 3.4.5.2 for detailed methods of FaME quality assurance. To ensure the highest possible standardisation of the FaME intervention across sites and amongst individual PSIs the following measures were achieved:

- All recruited instructors were in possession of their PSI Qualification.
- All PSIs attended one day's trial training.
- PSIs were issued trial session plans.
- Four standardised Quality Assurance (QA) visits were made to each FaME class at pre-set weeks in the intervention.
- Standardised QA paperwork was used to assess intervention fidelity and record action points for the PSI. (See Figure 3.3)
- Follow-up QA visits assessed achievement of previous PSI action points.
- Standardised home exercise booklets were issued to patients at a pre-set time point in the intervention.

### **3.2.10.2 OEP quality control**

Standardisation of the OEP intervention across sites and amongst individual PMs was challenging. Firstly, unlike the PSIs, the PMs were volunteers, so it was difficult to justify a strict QA protocol as this may have been a barrier to volunteering. Also, there were far more PM contacts with patients to observe than FaME classes (as patients in the OEP arm were seen individually), and the majority of contacts were by telephone, so for all these pragmatic reasons it was decided that direct observation of PM work was unsuitable. The following measures were achieved in an attempt to standardise the OEP intervention:

- 1) All PMs attended two days trial training
- 2) Standardised home exercise booklets were issued to patients at the start of the intervention
- 3) Four meetings were scheduled at pre-set weeks in the intervention for trial QA staff to meet groups of PMs
- 4) Trial paperwork (Patient Contact sheets) was issued to PMs to encourage standardisation in quality, volume and content of data collection from PMs regarding their contact with patients
- 5) QA staff reviewed PM paperwork for evidence of progression in exercise intensity.

### **3.2.10.3 Patient intervention evaluations**

The perceived quality of PSIs and PMs was also rated by the patients using written evaluations. I designed the evaluation documentation, which included questions to evaluate the intervention itself, as well as the patient's allocated PSI or PM. The data relating to the perceived quality of the PSI and the FaME intervention was used in my study and is therefore detailed later, in section 3.4.5.4.

### **3.3 Falls**

#### **Statement of IP**

The analysis of falls incidence in the intervention groups compared with the control was part of the ProAct65+ secondary outcome analysis plan. The intellectual property of the falls analysis therefore belongs to ProAct65+. However as falls was the main outcome for my PhD and of specific interest to me, I led on the falls analysis, under the supervision of Professor Richard Morris. I entered and cleaned this data, analysed it and wrote up the findings in a paper (Gawler, submitted for publication).

#### **3.3.1 Aims & objectives**

The falls analysis methods (and results) underpin my thesis as falls is the patient outcome under investigation in my therapist effect study. The methods relating to falls data are therefore described in detail in this section (3.3). The methods for my therapist effects study can be found in the section 3.4.

Falls rate reduction was a secondary outcome of the ProAct65+ trial. The hypothesis for this secondary outcome was that 24 week interventions, which were shorter in duration than in the previous research for both FaME and OEP, would reduce falls in the general population aged 65 and over; a younger population than in previous research for OEP, and at a lower falls risk than in previous research for FaME (see Chapter 1, sections 1.10.1 and 1.10.2).

#### **3.3.2 Falls outcomes & falls data collection**

**Falls outcomes** were:

- 1) Falls (per person) rate was compared in the interventions with the control (rate ratio).
- 2) Injurious falls (per person) rate was compared in the interventions with the control (rate ratio).

Although the full list of ProAct65+ outcome measures, including those relating to falls, has been shown in section 3.2.8, the falls-related outcome measures are discussed here in more detail; covering the purpose and protocol for each, as well published evidence regarding validity and reliability. Falls data collection time points are shown in Figure 3.1. Falls rate data was available for the full period of ProAct65+ (up to 24 months following the completion of the intervention), however, other falls outcomes were only measured at baseline and immediately post-intervention.

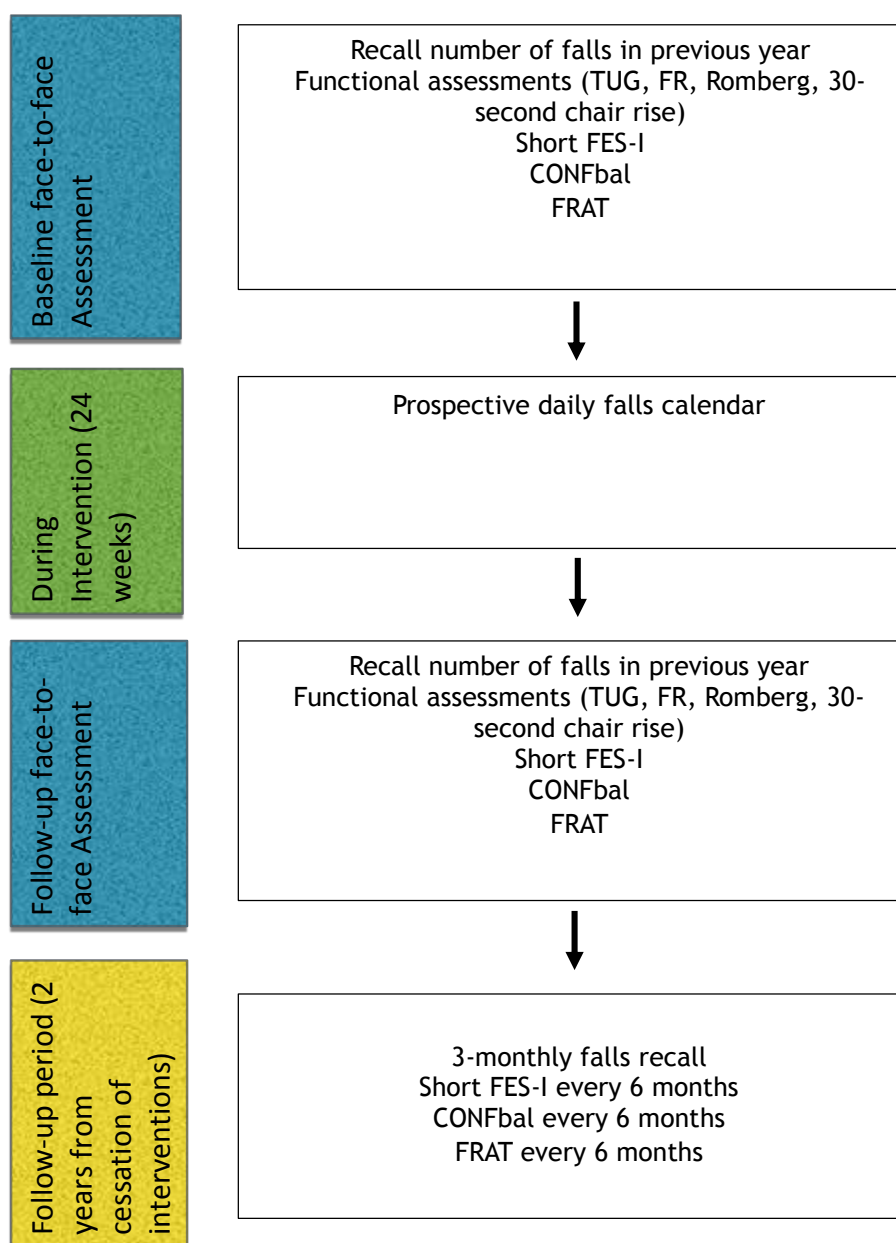
### **3.3.2.1 Falls numbers & falls diaries**

Falls reporting methods have been discussed in detail in Chapter 1, section 1.6. Daily, prospective reporting of falls is the 'gold standard' for falls data collection and is recommended in International guidelines for falls research (Hannan et al., 2010, Lamb et al., 2005). Hannan and colleagues reported validity as follows; sensitivity was 75%, and specificity was 96%, indicating that more fallers than non-fallers will be incorrectly identified. This can most likely be accounted for by the fact that some falls will not be captured even with prospective reporting. The reliability, analysed using the Kappa statistic, was 0.74 (95% CI 0.68, 0.80) (Hannan et al., 2010), however, since prospective reporting was considered the gold standard, this is a measure of how reliable retrospective reporting is, compared with prospective.

The number of fallers and falls in the year preceding the study were ascertained at baseline interview using a single question; "How many falls have you had in the last year?". During the 24-week intervention period patients were asked to complete a daily falls diary and return it in 4-weekly blocks (Figure 3.2). Those who failed to return their diaries received a reminder telephone call. Any inconclusive (poorly reported) falls and falls resulting in more serious injuries or hospitalisation were also followed up by telephone contact. At the follow up interview (immediately post intervention) patients were again verbally asked about their falls to act as a method for potentially infilling any missing falls diary data. During the two-year follow-up period, participants were asked to recall any falls over the preceding 3 months (rather than daily falls recording). This was a protocol amendment following high drop-out rates due to the

reported 'research burden' (the high number of questionnaires and diaries to complete) (Stevens et al., 2013).

**Figure 3.1,** Falls and falls risk factor data collection flow chart



**Figure 3.2,** Example ProAct65+ daily falls diary

		Mon	Tues	Wed	Thurs	Fri	Sat	Sun
FALLS DIARY								
	No fall	√	√	√		√	√	√
If you fell, what injury did you have, if any?	No injury				√			
	Bruise or cut							
	Muscle or ligament							
	Broken bone							

### 3.3.2.2 Functional assessments & falls risk assessment

Baseline functional assessment scores were compared to published normative data. Functional assessments were conducted as measures of balance and/or falls risk factors. All functional assessments were only conducted at baseline and immediately following the 24-week intervention. Self-reported measures were completed every 6 months up to 24 months post-intervention (see Figure 3.1).

#### 3.3.2.2.1 Timed-Up-and-Go

Timed-Up-and-Go (TUG) is widely-used and is recommended by the American Geriatrics Society/British Geriatrics Society as an assessment of falls risk. It was used in ProAct65+ as a measure of balance, leg power and falls risk, following the test protocol from the original TUG validation paper (Podsiadlo and Richardson, 1991). This study reported good validity (correlation coefficients of -0.81, -0.61, and -0.78, for Berg Balance Scale, gait speed and Barthel Index of ADL, respectively) and reliability (ICC 0.99) when TUG was used as an assessment of functional mobility. ProAct65+ patients were required to rise from a chair, walk a distance of 3 metres, turn around and walk back to the chair, then return to a seated position. The time to complete the task was

measured using a stop-watch and recorded, along with the need to use the arms for assistance when rising from the chair, and if a walking aid was used. Several studies have reported data regarding the ability of TUG to identify those at high risk of falls and the associated TUG time cut-off values for this have been varied (see Chapter 1, section 1.5). We used a TUG time of 13.5 seconds (Shumway-Cook et al., 2000). Shumway-Cook and colleagues reported good sensitivity (87%) and specificity (87%) when TUG was used to identify individuals who were at risk of falls, thus proving its validity. Moreover, Shumway-Cook recruited a similar population to the ProAct65+ population; community dwellers aged 65+, including fallers and non-fallers, and found an association between  $\geq 13.5$  seconds and retrospectively reported falls in the immediate pre-study period. The latter was of interest as ProAct65+ collected retrospectively reported falls data in the 12 months preceding the study.

However, owing to some speculation as to whether TUG alone is useful for discriminating fallers from non-fallers in a high-functioning, healthy older adult population (Schoene et al., 2013), ProAct65+ used the Falls Risk Assessment Tool (FRAT) (Nandy et al., 2004) as the primary method to identify those at high risk of future falls.

### **3.3.2.2.2 Falls Risk Assessment Tool**

The Falls Risk Assessment Tool (FRAT) is a 5-item questionnaire developed to identify those at high risk of a future fall in primary care (Nandy et al., 2004). The questions (see Chapter 1, section 1.5) were designed to be administered by a health or social care professional. A score of at least 3 identified those participants who would fall in the following 6 months with high specificity (0.92, 95% CI 0.88–0.94) but with low sensitivity (0.42, 95% CI 0.32–0.54) (Nandy et al., 2004). The low sensitivity indicates that the assessment identifies many false negatives; those who will fall, but are not identified as being at high risk. The low sensitivity did not present a problem for ProAct65+ as patients who fell 2 or more times in the year before expressing interest in the trial were excluded. FRAT was used as an outcome to monitor change over time,



as well as to dichotomise the ProAct65+ population according by falls risk as part of the baseline descriptive statistics.

#### **3.3.2.2.3 Functional Reach**

Functional Reach (FR) was used in ProAct65+ as a measure of balance and falls risk, following the test protocol from the FR validation paper (Duncan et al., 1990). This study examined the validity of FR as an assessment of stability, by comparison to a laboratory measure of stability, and the test-retest reliability across days. The authors reported good criterion validity (correlation coefficient = 0.71) and good reliability (intraclass correlation coefficient (ICC) = 0.81). ProAct65+ patients were required to stand and raise one arm at shoulder height (start position), then reach forwards as far as possible (finish position). The distance from the start position and the furthest point reached was recorded. ProAct65+ used a FR score of 15cm (Duncan et al., 1992) as a cut-off point to identify fallers as part of the baseline descriptive statistics. Predictive validity reported by Duncan (1992) was the Odds Ratio of having 2 or more falls in the 6 month follow up period was 4.02 (95% CI 1.84-8.77) for those with a reach of 6 inches (15.24cm) or less, compared to those with a greater reach.

#### **3.3.2.2.4 Modified Clinical Romberg Static Balance test**

Modified Clinical Romberg Static Balance test, eyes open and closed (Freeman, 1965) was used as a measure of balance. Poor balance is a known risk factor for falls (Lord et al., 1992, Todd and Skelton, 2004). The protocol from the FICSIT static balance measures validation paper (Rossiter-Fornoff et al., 1995) was used in ProAct65+. This incorporates the original Romberg position (Freeman, 1965). Rossiter-Fornoff (1995) studied test-retest reliability and construct validity, reporting a correlation coefficient of 0.66, over 3-4 months, for the former. However, such a long period between test and retest, even in control subjects, assumes that balance in older subjects does not deteriorate over time. The validity was not evaluated against a 'gold standard' balance measure, and appears low to moderate with correlations ranging from 0.10 to 0.52. However, the Romberg assessment was used in ProAct65+ alongside other measures of balance.

In this assessment, as per the protocol (Rossiter-Fornoff et al., 1995), the patient was required to adopt progressively more challenging stances and to 'balance' without the use of the upper limb for 10 seconds. The level of challenge of each stance is determined by the base of support, which is reduced in each subsequent stance, for example; a semi-tandem position is followed by a tandem position. Each position (except the single leg stand) was also attempted with the eyes closed, should the patient successfully perform 10 seconds with the eyes open. The patient is assigned a total score for the test, the maximum being 28. The Romberg Balance test requires no specialist equipment and is quick to administer. There is no published cut-off point for identifying fallers using this tool.

#### **3.3.2.2.5 30 Second Chair Rise**

30 second chair rise was used in ProAct65+ as a measure of lower limb strength and power. Poor leg power is a known risk factor for falls (Lord et al., 1992, Todd and Skelton, 2004). The test protocol from the 30-second chair rise validation paper (Jones et al., 1999) was followed. Jones (1999) reported good test–retest reliability of 0.89 over two days and good validity of 0.87 when 30 second chair rise was compared to a gold standard criterion test.

The assessment was performed as follows; the patient starts seated (with the chair back against a wall to prevent the chair from moving) and performs as many full sit to stands as is possible in 30 seconds, with the arms folded across the chest. The number of stands achieved was recorded.

#### **3.3.2.3 Fear of falling**

Falls Efficacy Scale-International and Confidence in Maintaining Balance were used as measures of fear of falling and confidence in maintaining balance during everyday tasks.

#### **Falls Efficacy Scale-International**

Falls Efficacy Scale-International (FES-I) is a reliable and valid measure of concern regarding falling whilst participating in a range of activities of daily living (Yardley et al.,

2005). ProAct65+ used the shorter 7-item version which is suitable for older people with different levels of concern regarding falling (Delbaere et al., 2010). Yardley and colleagues (2005) reported excellent test–retest reliability (ICC=0.96) and their concurrent validity analyses demonstrated higher FES-I scores for subjects with a history of falls and/or risk factors for falling. A higher total score indicates poorer self-efficacy, with a maximum possible total score of 28 for the 7-item (Short) FES-I. The published cut off point of 11, which differentiates between low and high concern about falling (Delbaere et al., 2010), was used to dichotomise baseline Short FES-I scores. Delbaere (2010) reported excellent convergent and predictive validity; 100% and 93%, respectively.

### **Confidence in Maintaining Balance**

Confidence in Maintaining Balance (CONFbal) is a validated and reliable measure of confidence to perform a range of activities of daily living (such as climbing stairs) without falling (Simpson et al., 1998). Simpson (1998) investigated test-retest reliability a week apart and reported excellent reliability (ICC=0.96). Construct validity was also studied by investigating correlation with a range of other measures; with 0.51 to 0.74 reported, demonstrating good validity. A higher total CONFbal score indicates poorer confidence, with a maximum possible total score of 30.

### **3.3.3 Data management**

I entered the falls data into an SPSS falls database. Double data entry was carried out for 1 in 10 records by a second researcher. Diary data were sometimes difficult to interpret, usually because the patient had used the diary incorrectly, for example, ticking more than one box per day. Diary data was only used in the falls analyses if the patient gave an unequivocal report of a fall. The impact of this on findings will be discussed in Chapter 6, section 6.4. All researchers entering diary data used the same procedure (from a Standard Operating Procedure that was written to ensure

standardisation of diary data entry) to decide if the data were interpretable.

### **Protocol violations**

Participants who reported more than 2 falls in the year preceding the study (but who had not been excluded by the researcher at baseline) were deemed to be protocol violators and were removed from the falls analysis.

### **Missing data**

Missing falls diary data was accounted for by calculating a time at risk (of falls) for each patient based on the number of diaries they completed, for example, if all 6 diaries were completed and indicated 2 falls, 2 falls in 24 weeks (at risk) was entered, whereas if only 2 diaries were completed and indicated 2 falls, 2 falls in 8 weeks (at risk) was entered. This method assumes the rate of falls would be constant over time for each patient. This will be discussed later, in Chapter 6, section 6.4.

### **3.3.4 Baseline falls and falls risk analysis**

The proportion of patients having reported a fall in the year preceding the study (fallers) in the ProAct65+ population was compared to the proportion of fallers within the older UK population. Numbers of falls and fallers were also compared in the three groups (FaME, OEP and usual care) at baseline. The proportion of participants 'at risk' of falling was compared across trial groups. The primary method of identifying those at high risk of falls was using the Falls Risk Assessment Tool (FRAT). For descriptive characterisation of the patients (at baseline only) we also dichotomised them according to falls risk status using Timed-Up-and-Go (TUG) and Functional Reach (FR). Other risk factors for falling, including poor balance and fear of falling, were also described at baseline. The rationale for outcome measure cut-off points was described

earlier, in section 3.3.2.2. The methods used to identify those at risk and to describe the recruited population in terms of falls risk factors at baseline were as follows:

- 1) Scoring  $\geq 3$  (out of a maximum score of 5) using the FRAT
- 2) Taking longer than 13.5 seconds to complete the TUG test
- 3) Having a FR of less than 15cm
- 4) Short FES-I score  $\geq 11$

### **3.3.5 Falls analysis**

Falls rates were calculated by dividing the number of reported falls, in each time period for each patient, by the time at risk of falls. The time at risk of falls was indicated by the number of diaries returned. Falls data were entered into SPSS (version 21) and cross- tabulations were undertaken, stratifying falls by trial time-point and site. Falls data were analysed using negative binomial modelling on an intention to treat basis accounting for clustering by practice. Outputs from the negative binomial analysis were in the form of Incidence Rate Ratio (IRR), 95% CI and p-values. Negative binomial modelling is recommended for falls prevention programme evaluation (Robertson et al.,2005). Other statistical approaches that are also potentially suitable for analysing the efficacy of falls prevention interventions due to their ability to allow for the non-normal distribution of falls and for varied patient follow-up times, have been shown to violate the underlying assumption of proportional hazards (Robertson et al., 2005). This assumption is described as “the ratio of the risks of the events in the two groups is constant over time” (Robertson et al., 2005). In relation to falls analysis, this means that the ratio of the risk of falling in the intervention group and the control group remains the same over the trial duration.

Falls rate was compared between each intervention group and the control during the intervention period and each post-intervention year separately, as well as for the combined intervention period and first post-intervention year (18 months in total). A sensitivity analysis was carried out to see if diary data were missing at random across study arms, and to investigate if any patient characteristics (gender, age, falls rate, number of co-morbidities) were associated with diary returns rate. A per protocol analysis was additionally carried out comparing falls incidence rates between only

those in each intervention arm who adhered to at least 75% of the exercise programme with the control group. An additional per protocol, within group analysis, comparing those in the FaME arm who continued to achieve 150 minutes of MVPA per week into the second post-intervention year compared to those in the FaME group who did not maintain their physical activity, was added to the planned ProAct65+ falls analyses.

### **3.4 The therapist and dose effects study**

#### **3.4.1 Aims**

The aims of this research were to study (a) the effect of the therapist (delivering the group exercise programme), and (b) the effect of exercise dose, on falls incidence and falls risk factors within a 24-week group exercise programme (FaME). Study objectives and hypotheses are detailed in Chapter 1, section 1.13.

#### **3.4.2 Patient selection**

In line with my study aims, my study participants were the ProAct65+ FaME group patients. However, it was not possible to study the effect of the therapist on those FaME subjects who were not allocated a therapist. Reasons for FaME subjects not being allocated a therapist included GP exclusion and withdrawing at allocation. An intention-to-treat approach could not be adopted since no predictor variables were available for those patients who were not allocated a therapist. Patient selection for my study was therefore limited to those people in the ProAct65+ FaME group who were allocated a therapist, even if they did not attend any of the exercise classes offered.

#### **3.4.3 Sample size**

As identified in section 3.2, the ProAct65+ trial was not powered to detect effects of the interventions on falls. As my study was nested in the ProAct65+ trial, the number of participants in my study was dictated by the ProAct65+ sample size calculation to detect effects of the interventions on physical activity. A post-hoc sample size calculation for a bespoke evaluation of therapist effects on falls outcomes has been

presented in Chapter 5, section 5.5, in order to assess whether the ProAct65+ sample would have been adequate to detect a difference in the main falls outcome in my study.

#### **3.4.4 Falls & falls risk factors outcome measures**

For details regarding falls data collection see section 3.3.2.1. Risk factors for falls include poor lower limb power and balance (Chapter 1, section 1.4). Lower limb power and balance were investigated at baseline and immediately post-intervention using the four functional assessments (see section 3.3.2.2). Any change in functional assessment score from baseline to the end of the intervention, therefore, indicated a change in falls risk factors.

#### **3.4.5 Data collection and management**

##### **3.4.5.1 Postural Stability Instructor (PSI) variables**

PSI variables, also referred to as PSI characteristics in this thesis, included age, gender, professional background, previous experience of delivering the FaME intervention, timing of attainment of PSI qualification (before or as part of the trial), the ‘quality’ of intervention delivered, the patient-rated intensity of intervention delivered, patients’ enjoyment of intervention delivered and attendance (proportion of the intervention the allocated PSI taught themselves). PSI variable data were collected from various source documents including the PSI’s CV, their Criminal Record Bureau (CRB) paperwork, the class attendance registers and from patient evaluations of the intervention. The source data are listed in Table 3.4. Any calculation of rating or score from source documents (for example PSI quality rating from quality assurance documentation) has also been indicated in Table 3.4 and is covered in detail in section 3.4.5.3.

##### **3.4.5.2 FaME Quality Assurance (QA) visits**

QA visits were made in weeks 2, 8, 16 and 24. I devised the QA paperwork based on the Later Life Training PSI Summative Assessment Checklist, used in the PSI’s practical observed assessment which forms part of the PSI qualification exam. The original document was designed by a team of specialists (physiotherapists, researchers,

exercise physiologists) working with frailer, older adults and includes assessment criteria relating to the skills needed to deliver the FaME falls prevention intervention. The checklist I developed included these delivery/teaching criteria along with some additional trial-specific criteria such as ‘submitted completed attendance register on time’ and ‘patients submitting diary data at levels similar to other classes’ (Figure 3.3). The checklist was used by the person assessing the QA (myself or one other colleague) in conjunction with a feedback sheet on which were recorded field notes relating to the QA criteria.

**Table 3.4,** PSI variables & data sources

Variable	Data Source	Calculation Method
PSI Gender	CRB	-
PSI Experience	PSI CV	-
PSI Training	PSI CV	-
PSI Professional Background	PSI CV	-
PSI Attendance	FaME attendance registers	Percentage of sessions taught by allocated PSI (not cover teacher)
PSI Quality	QA Checklist	See section 3.4.5.3
PSI Age	CRB	Age during year of intervention delivery
Patient attendance (dose)	FaME attendance registers	Percentage attendance of number of sessions offered
Patient evaluation question ‘were the classes the appropriate intensity for you?’	Patient evaluation	-
Patient evaluation question ‘did you enjoy the classes?’	Patient evaluation	-

Each field note or observation was linked to a criterion by writing the criterion number next to the comment on the feedback sheet. The QA staff member would also record a tick (to indicate the item was achieved), a star/asterisk (to record a comment), a Q (to record a question) or an S (to record an action that needed support) in the box adjacent to the selected criterion. There is a key in the top left of the QA Checklist showing the range of options available to the QA assessor. This type of checklist is



useful in the standardisation of exercise delivery as it focuses the person assessing the QA on pre-determined teaching skills and therefore standardises the quality assurance process itself. This was important in ProAct65+ where more than one researcher was making QA visits. In addition to this, the first few QA visits were attended by both researchers assessing the QA to further ensure standardisation of the QA process. PSIs were supplied with the QA paperwork prior to commencing intervention delivery so they too were aware in advance of the trial requirements. PSIs were scheduled to meet with QA staff immediately after QA observation of FaME sessions so that feedback, action points and any associated support could be discussed and planned. Copies of QA paperwork were stored by the trial and the original documentation was given to the PSI. Action points were reassessed at subsequent QA visits.

QA documentation was used to assign a quality 'rating' to each PSI for the therapist effect analysis. The method of calculating the PSI quality rating is reported in section 3.4.5.3. The rating was entered into the SPSS database along with other PSI 'characteristics' such as gender, age and professional background.

**Figure 3.3, PSI Quality Assurance Checklist**

PSI Name		QA visit number	QA visit date	Venue					Name of observer		
<b>Preparing</b>		<b>Teaching</b>		<b>Warmup</b>	<b>Dynamic endurance</b>	<b>Dynamic Balance</b>	<b>Seated /Standing Resistance</b>	<b>Backward Chainin g</b>	<b>Floor work</b>	<b>Cool-down Flexibilit y</b>	<b>Cool-down Adapte d Tai Chi</b>
Arrived in time to meet participants		Engaged participants in order to motivate and promote confidence									
Completed safety check on venue		Selected safe and effective exercises appropriate to the component.									
Wore attire appropriate to the activity		Selected safe and effective exercises appropriate to the stage in the intervention									
Appropriately arranged the group, individuals and resources		Selected the appropriate speed for the exercises									
Welcomed participants		Gave effective visual and verbal instructions									
Took register of attendance		Provided specific relevant teaching points to enhance technique, effectiveness and postural stability									
Verbally screened participants for falls, previously reported injuries and new or known medical conditions		Reinforced the specific relevant teaching points at regular intervals									
Appropriately followed up returners after period of absence		Provided safe transitions between exercises and session components									
Reminded attenders to keep up with and submit diaries		Demonstrated and performed exercises									

	accurately and with good posture								
Ensure that infection control procedures are implemented and adhered to	Changed teaching position to improve observation and enhance communication								
Ensure that confidentiality of personal and medical data is respected	Demonstrated the use of observation and effective correction								
<b>Liaison with research team</b>	Explained the purpose of the exercises, relating them to postural stability and daily life								
Submitted completed register on time	Encouraged interactive communication, to check or clarify understanding, with group and one to one.								
Evidence of telephone follow-up of non-attenders	Spoke clearly, audibly and at an appropriate pace								
Patients submitting diary data at levels similar to other classes	Adapted exercises to meet the needs of participants with postural stability challenges								
This class attendance data similar to other classes?	Offered alternatives to allow for different levels of ability / tailored exercises to individuals								

#### **3.4.5.3 Use of quality control data to calculate therapist 'quality' categorisation**

As described above the QA checklist was filled in by the QA researcher at each site visit. The QA researcher inserted a tick, an asterisk, a Q or an S into each box adjacent to each performance criterion. This is a system used in the training and appraisal of exercise professionals and as such was already familiar to several of the trial PSIs. Those trial PSIs from a non-exercise background (the physiotherapists) were made familiar with the process at the trial top-up training. For the purpose of the therapist effect (TE) analysis, the checklist was used to 'calculate' a therapist 'quality' categorisation; low, moderate or high. This was a novel use of the checklist devised specifically by me for the TE analysis. Each symbol (tick, asterisk, S or Q) on the QA checklist was given a numerical score as follows; 1 for a tick, 1 for any positive comment linked to an asterisk, -1 for any actions (S) and -1 for any negative comment linked to an asterisk. A total score for each checklist was thus calculated, then a total score for each PSI was achieved by adding all checklist scores for each PSI, divided by the number of visits they had received i.e. a mean checklist score for each PSI was calculated. The highest possible score was 143.

Mean scores of between 0 and 80 were assigned the low quality categorisation, those of between 81 and 119 were assigned the moderate quality categorisation, and those of 120 or more were assigned the high quality categorisation. These cut off points were determined according to the anticipated spread of scores. Theoretically, the lowest possible score was zero, but every box on the checklist would need to have contained an S, or an asterisk linked to a 'negative' comment, for zero to be allocated. In reality, this would not happen, as the assessor would be mindful of ensuring a balance of positive and 'negative' comments. The lowest score was therefore more likely to have been in the region of 50/60. Potential mean scores ranging from 55 to 143 were divided into tertiles; 55-84, 85-113, 114-143. This process was undertaken by myself and independently by a second researcher who had not been involved in the QA of trial PSIs to ensure 1) the categorisation by the agreed process was accurate and 2) to reduce bias in the interpretation of comments as being either positive or negative. To ensure that the calculation and assignment of quality process resulted in the most appropriate quality 'rating' for each PSI, the two QA researchers

independently rated each PSI either low, moderate or high without the use of the checklists and this was compared to the calculation method. Any discrepancy between the subjective view of each assessor and the calculation (based on the subjective view of the assessor) was discussed and a consensus decision regarding the final 'rating' was reached.

#### **3.4.5.4 Patient evaluations**

The Patient evaluation was a self-complete questionnaire issued by post immediately following the 24-week intervention period. I devised this questionnaire and included questions that related to the perceived quality of the exercise instructor (PSI) and the perceived quality and benefit of the exercise intervention. In particular, the questions were trying to measure patients' enjoyment of the exercise, their perception of the intensity of the exercise for them, their feelings about the venue in which the exercise was undertaken and the refreshments offered after the exercise, and their opinion regarding their instructor's professionalism, for example, starting and ending sessions on time. The questionnaire was designed with both the therapist effect analysis in mind and as an objective of ProAct65+ to gather data relating to patients' perception of the value of the intervention. For this reason, the questions focused on issues broader than therapist quality. However, the aspects of therapist quality that I measured using the evaluation questionnaire were 1) the ability of the PSI to tailor the intensity of the exercises to the individual, 2) the ability of the PSI to make the exercise enjoyable for the individual (possibly by tailoring, and/or by displaying aspects of professionalism and/or by having rapport with individuals) and 3) the professionalism of the PSI in terms of running the intervention (for example, arriving in time to greet people). The data from this evaluation was entered into the SPSS database along with other 'therapist' characteristics. See Appendix 3 for the patient evaluation form.

#### **3.4.5.5 Patient attendance (dose)**

All patients in the FaME group were offered 24 weeks of an hour-long once-weekly supervised exercise class, therefore, the dose of supervised exercise under investigation in the main trial (ProAct65+) was 24 hours. Attendance at the weekly exercise class was recorded by the PSIs on attendance registers. I entered all attendance data from the class registers into an SPSS database and calculated an actual total volume of supervised exercise (in hours) achieved by each patient, as well as a percentage (of the total intervention) that each patient achieved. These data are referred to as the 'dose' of exercise (per patient) in my study.

#### **3.4.5.6 Data entry**

I designed and set up the SPSS 'therapist/dose' database and entered all data. I cleaned the data by double-entering all data, then checking for discrepancies and correcting by returning to source documents (source documents are listed in Table 3.3). Further cleaning was achieved via random 1 in 10 checks that were carried out by another researcher (ZS). In the case of any disagreement, source documents were re-checked. Table 3.4 indicates which trial documentation was used to populate the SPSS variables within the 'therapist/dose' database.

#### **3.4.6 Blinding**

I was not blinded to the therapist allocation because I also evaluated (by QA visits) the therapists and assessed patients (who were inclined to mention the name of their exercise class teacher) within the trial. Furthermore, I was not blinded to patient dose as I entered data from class attendance registers, although it would be difficult to remember attendance data relating to so many patients. The lack of blinding in these areas was not planned as neither dose nor therapist effects were outcomes of the ProAct65+ trial.

### **3.4.7 Therapist and dose effects analysis plan**

#### **3.4.7.1 Data cleaning**

Histograms were produced for outcome variables to assess the normality of distribution. The frequencies of all variables were tabulated to check for outlying values.

#### **3.4.7.2 Baseline characterisation**

Descriptive analysis of the baseline characteristics of the study therapists was undertaken, in the form of means and percentages. Descriptive analysis of the ProAct65+ participants, stratified by randomisation group, was undertaken as part of the ProAct65+ falls analyses, and therefore the associated methods have already been described in section 3.3.4 of this thesis. The therapist and dose effects study participants were a sub-set of the ProAct65+ FaME group, so descriptive analysis of the baseline characteristics of this group was undertaken, as was a crude comparison with those patients from the FaME group who were ineligible for inclusion in the therapist effect group.

#### **3.4.7.3 Bivariate analyses**

The dependent variables were falls rate and falls risk factors. The predictor variables were therapist characteristics. A total of 7 dependent variables (DVs) were used; 4 functional assessments (Timed-Up-and-Go, Romberg balance, 30-second chair rise & Functional Reach) from the follow-up meeting with patients immediately post-intervention (which provide data relating to falls risk factors such as leg power and balance) and falls rates from 3 trial time periods; 1) during the 24 week intervention, 2) during the 12 months following the intervention, and 3) during the 24 week intervention and the 12 months following the intervention (a total of 18 months from baseline). The three falls rate periods were chosen because the original FaME trial (Skelton et al., 2005) had shown a non-significant increase in falls during the intervention period for the exercise group compared with the control group, and this was thought to be due to the increased exposure to falls risk that the intervention itself posed for the frequently-falling population the study had recruited. In the ProAct65+ falls analysis, we therefore wanted to investigate the immediate effect of

the intervention on the ProAct65+ population separately to the post-intervention period in the ProAct65+ falls analysis. As my study was an investigation into the effect of the therapist on falls rate, it was logical to follow the ProAct65+ model and study the intervention period separately to the post-intervention period to see if the therapist had an effect on the exposure to falls risk during the intervention, and separately to see if, in the longer term (post-intervention), the therapist had an influence on the overall protective effect of the intervention against falls. Falls data from the second post-intervention year were not used in my study as we had reported in the ProAct65+ HTA report that the effect of the FaME intervention on falls incidence was lost in the second post-intervention year.

The ten predictor variables in my therapist effect analysis included therapist age, gender, professional background, clinical experience, timing of PSI training, reliability and 'quality', as well as the dose of the intervention each patient achieved, and patient responses to two evaluation questions (enjoyment and intensity). Although intervention dose may appear to be a strictly patient-level factor, rather than a therapist-related factor, good adherence to instructor-led exercise is often associated with positive rapport between instructor and patient, therefore dose was explored in this study as a potential indicator of therapist 'skill'. The evaluation responses were analysed for much the same rationale; patient enjoyment of the exercise sessions may imply that the instructor was skilled in the delivery of the exercise intervention (for example, by ensuring the patient was comfortable whilst exercising by offering appropriate adaptations) and patient evaluation of the intensity of the exercises implies therapist skill at appropriately selecting the correct level of challenge for each individual within the group.

Bivariate analyses were conducted for each predictor variable with each dependent variable separately, to explore potential relationships between the pairs of variables and to inform the more complex analyses. Correlation, T-test or Analysis of Variance (ANOVA) was selected depending on the type of variables within the pair (categorical variable versus continuous variable, for example). All planned bivariate analyses are shown in Table 3.5. Correlation analysis was used to explore the strength and direction



of linear relationships between the paired variables. Outputs from the correlation analyses are in the form of scatter plots, correlation coefficients and p-values. T-test analysis was used to compare the difference in group means (for example, mean falls rate in patients taught by male PSIs compared with that in patients taught by female PSIs) and outputs are in the form of t-statistics (df) and p-values. ANOVA also compares group means and outputs are in the form of F-statistics (df) and p-values. The ANOVA analyses are accompanied by post-hoc tests (Tukey) in order to investigate which group means differed, when the ANOVA output showed an overall difference between groups.

Prior to undertaking the analyses described in Table 3.5, mean falls rate and standard deviations were tabulated, stratified by therapist characteristic. Mean post-intervention functional assessment scores and standard deviations were also tabulated, stratified by therapist characteristic.

**Table 3.5, Bivariate analysis selection**

Paired variables	Analysis
PSI age, intervention dose & PSI attendance separately, with falls rate, TUG score, functional reach score, romberg balance score, 30-second chair rise score.	Pearson bivariate correlation with 2-tailed significance test
PSI gender, professional background, clinical experience & training separately, with falls rate, TUG score, functional reach score, romberg balance score, 30-second chair rise score.	Independent samples T-test
‘Quality’ of therapist, patient enjoyment of classes & patient-rated intensity of classes separately, with falls rate, TUG score, functional reach score, romberg balance score, 30-second chair rise score.	ANOVA one way Tukey

#### **3.4.7.4 Multiple regression equations**

Multiple regression equations (for falls and falls risk factors separately) were planned to identify independent effects of each therapist characteristic that had a positive effect on falls rate or falls risk factors, adjusted for all others (from the crude bivariate analyses discussed earlier, in section 3.4.7.3). This was only considered necessary if

more than one therapist characteristic appeared to be associated with a patient outcome.

#### **3.4.7.5 Multilevel modelling adjusting for baseline clustering**

Outcome variables (falls rate and falls risk factors) and Postural Stability Instructor (PSI) level variables have been discussed already (section 3.4.4). No patient level variables were used in the multilevel modelling analysis other than the outcome variables. This was in line with the objective of my study to focus on any therapist effects on patient outcome variables.

PSI level data and outcome data were entered into one SPSS database. Outcome variables were entered by patient, as were PSI level variables. The data in this database therefore contained a two level structure; patients (at level 1) clustered by PSI (at level 2). They were analysed using multilevel modelling in SPSS version 21. Multilevel modelling of this hierarchical data allowed the evaluation of between-PSI level groups variability and between-patient variability in the falls rate and falls risk factor outcomes at the same time. The analysis method assigned the total variation in the falls rate/falls risk factors into between-PSI and within-PSI group parts. The first analyses were unconditional models (models without predictor variables) from which therapist effects on patient outcomes were determined by calculating the intraclass correlation (ICC). The ICC is defined as “a ratio of between [PSI] variability/(between [PSI] variability + within [PSI] variability)” (Li et al., 2005) and describes how strongly patients grouped by PSI resemble each other.

In this study ICC was used to provide a justification for the use of multilevel modelling. It was planned to only introduce predictor variables to the analysis if there was some evidence of a therapist effect provided by the ICC. In order to adjust for baseline cluster by GP practice, baseline functional assessment scores were added in to the unconditional models before any predictor variables were investigated. Predictor variables were added separately to allow the evaluation of their effect on the between-PSI level groups variability and on the between-patient variability in isolation,

thus allowing the effect of individual therapist characteristics to be investigated without confounding.

#### **3.4.7.6 Multiple testing**

One strategy to deal with significant p-values that could have arisen by chance from multiple testing is the resampling methods approach (Westfall and Young, 1993). It was planned to apply this method in the event of obtaining any results significant at the 5% level.

### **3.5 Chapter conclusions**

This chapter has described:

- 1) ProAct65+ methods that relate to my study.
- 2) ProAct65+ methods that relate specifically to falls data and the falls analysis.
- 3) The methods of my study into therapist and dose effects.

The third part of this chapter has described a novel, exploratory study developed to evaluate the effects of therapists and dose on falls and falls risk factor outcomes, following a falls prevention exercise intervention for general older people. The primary objective of the study was to establish the effects of the therapist and dose on falls over the intervention and the follow up periods. The secondary objective was to establish the effects of the therapist and dose on specific risk factors for falls; balance and lower limb power, over the intervention period.

## **Chapter 4**

### **Results 1; ProAct65+ baseline characteristics & ProAct65+ falls analyses**

#### **4.1 Chapter summary**

This is the first of two chapters reporting results. In the first part of this chapter I describe the baseline characteristics of the whole ProAct65+ population, from which my therapist effect group was derived. (The therapist effect group is described Chapter 5, section 5.3.) In the second part of this chapter I report the findings from the ProAct65+ falls analysis (on which I led). The falls results are essential foundation information for the second results chapter (Chapter 5, in which I report findings from my therapist and dose effect analyses) as falls and falls risk factors were the patient outcome variables used in my study. My study aims are described in Chapter 1, section 1.13.

#### **4.2 ProAct65+ recruitment**

1256 patients were recruited from 43 GP practices in London and Nottingham. 387, 411 and 458 were randomised into the FaME, OEP and UC arms, respectively. The required number of exercise instructors (PSIs) were recruited and trained in each site. A total of 12 PSIs were deployed, 6 in London and 6 in Nottingham. Fuller details of recruitment of patients and staff are published elsewhere ([www.ncbi.nlm.nih.gov/pubmed/25098959](http://www.ncbi.nlm.nih.gov/pubmed/25098959)). PSIs (referred to as therapists in my therapist effect study) are described in Chapter 5, section 5.2.

#### **4.3 Baseline ProAct65+ patient demographic characteristics**

The patient age range was 65 to 94 years (average age 73) with 84% of patients in the 65-79 age group. 779 patients (62%) were female and 176 (14%) were non-white. The mean number of co-morbidities and medications per patient was 1.7 and 3.7, respectively. In terms of baseline self-reported physical activity (PA) habits, 6% reported no PA, and the median number of minutes of moderate or greater intensity physical activity (MVPA) per week was 105. The proportion of patients already

achieving the Department of Health recommended 150 minutes of MVPA per week (DoH, 2011) at baseline was 40%, 41% and 38% in the FaME, OEP and UC groups, respectively. This compares favourably with Health Survey for England data regarding national PA levels; 20% men and 17% women aged 65 to 74 were reported to achieve 150 minutes of moderate intensity PA per week (HSE, 2009), but indicates that the screening of patients against the exclusion criterion relating to PA (see Chapter 3, section 3.2.5.1) prior to baseline assessment and randomisation was ineffective. Despite the ProAct65+ population reporting higher than average levels of PA, they had poorer than published average scores for most of the functional assessments (see section 4.5).

#### **4.4 Baseline ProAct65+ patient falls characteristics**

##### **4.4.1 Protocol violations**

Data from 18 patients were excluded from the falls analysis as they reported more than 2 falls in the year preceding the study. A patient who reported 76 falls during the intervention period, despite not reporting any falls in the year prior to the study, was also excluded from the falls analysis following a telephone conversation with him in which he revealed he had withheld information regarding his previous falls. We checked to see if any other patients had reported dramatically different numbers of falls during the study compared with prior to the study, but there were no other such cases. Two further patients withdrew from the study and requested removal of their data from the analyses.

##### **4.4.2 Baseline falls characteristics**

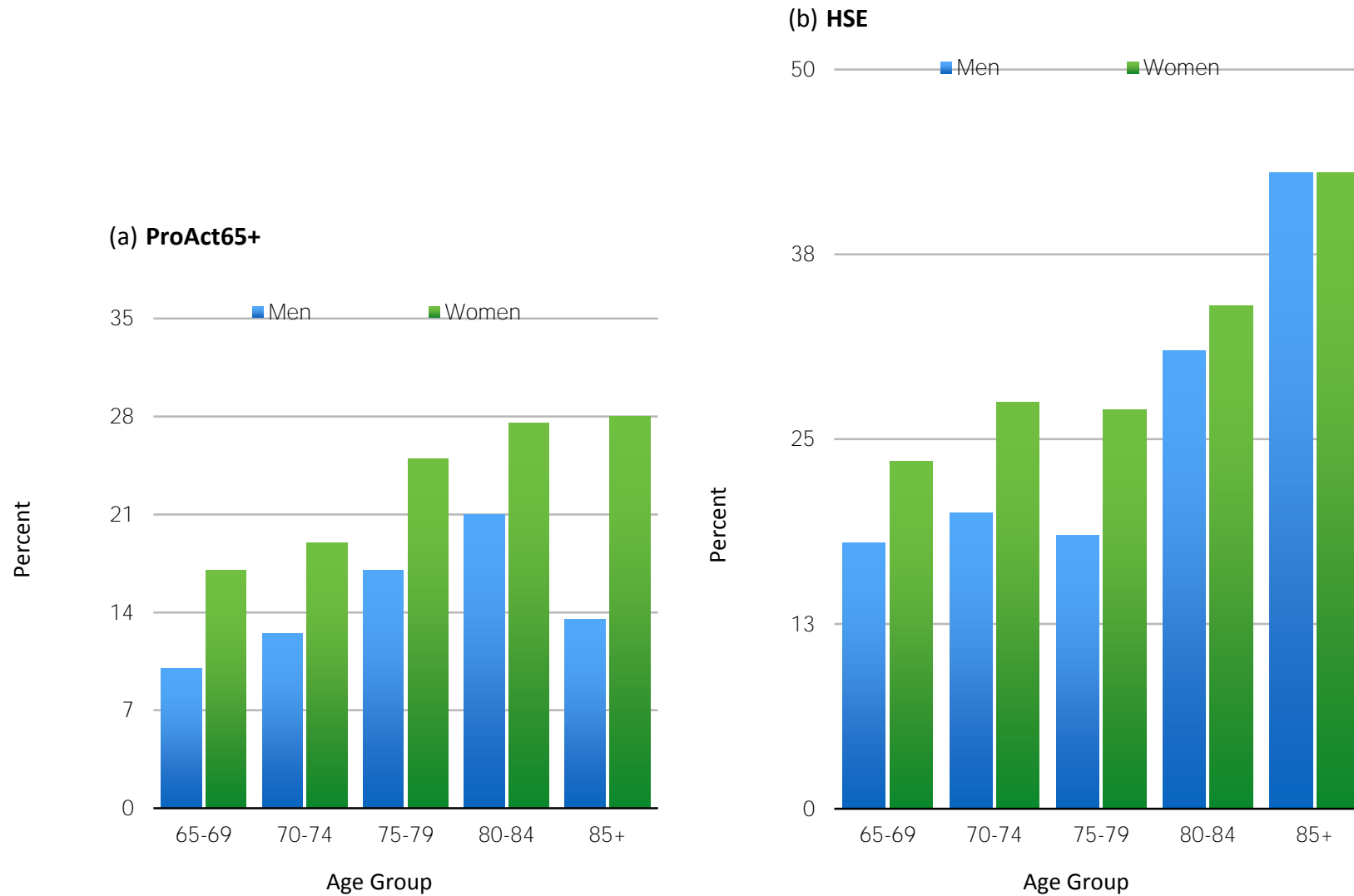
Of the 1235 patients included in the falls analysis, 294 patients (24%) reported 1 or 2 falls in the previous year (21% of men and 27% of women). Use of the Falls Risk Assessment Tool (FRAT) identified 76 (6%) patients as being at high risk of a future fall, 182 (15%) took longer than 13.5 seconds to complete the TUG test indicating a high risk of a future fall, 97 (8%) scored less than 15cm on the functional reach assessment

indicating a high risk of a future fall, and 209 (17%) scored  $\geq 11$  for falls self-efficacy indicating a high concern about falls.

#### **4.4.3 ProAct65+ population falls compared with UK older population**

One in 3 people aged 65 and over fall each year (Gillespie et al., 2012, Nyman and Victor, 2012, Tinetti and Kumar, 2010). Frequent fallers were excluded from ProAct65+, therefore the number of fallers would be expected to be lower than 1 in 3. Almost 1 in 4 (24%) of the ProAct65+ population fell in the year prior to the study. Being female and being older are both considered to be risk factors for falling (Close, 2005b). Indeed, 1 in 2 people aged 85 and over experience falls (O'Loughlin et al., 1993). Falls reported by ProAct65+ patients in the year preceding the study were analysed by gender and age (Figure 4.1(a)). In both men and women, the proportion reporting falls increased by age, up to age 84. In the 85+ group, however, whilst the percentage of women who fell continued to rise, the percentage of male fallers sharply declined. When the proportion of fallers in the ProAct65+ population was compared to Health Survey for England data (HSE, 2005) (Figure 4.1(b)) and 95% confidence intervals were applied, we found that the proportion in all age groups of both men and women up to age 84 was not significantly less than in the broader UK population. However, the proportion of fallers in the 85+ group of both men and women was significantly less than in the HSE sample (men, 95% CI -0.417, -0.183; women, 95% CI -0.281, -0.019).

**Figure 4.1,** Percentages with at least one fall in the year preceding the study, by age & gender, compared with HSE.



A further analysis (Pearson's chi-square test) of gender and age associations with falling revealed that, in the ProAct65+ population, female gender and older age were significantly associated with falls (Table 4.1).

**Table 4.1,** Falls reported in previous year, by gender and age, n (%)

	Totals	Reported 1+ falls	Did not report any falls	p value
Male	465 (38)	95 (20)	370 (80)	0.021
Female	768 (62)	200 (26)	568 (74)	
65 to 75 years	857 (70)	178 (21)	679 (79)	<0.001
76 to 94 years	376 (30)	117 (31)	259 (69)	

#### 4.4.4 Baseline falls characteristics by study group

At baseline, there were similar proportions of fallers (having had at least one fall) in all trial arms; 82 (22%) in FaME, 94 (23%) in OEP and 118 (26%) in UC. The average number of falls per person reported in the year prior to the study in each group was 0.27, 0.29 and 0.31 in FaME, OEP and UC, respectively. Further comparisons are shown in Table 4.2.



**Table 4.2,** Participants' baseline falls characteristics by group, n (%)

Outcome measure	FaME n=377	OEP n=404	Usual Care n=454	All n=1235
Reported fall(s) in previous year	82 (22)	94 (23)	118 (26)	294 (24)
Falls Risk Assessment Tool $\geq 3$ (higher score = greater risk of falls)	19 (5)	20 (5)	37 (8)	76 (6)
Timed Up and Go >13.5 seconds (higher score = greater risk of falls)	56 (15)	53 (13)	73 (16)	182 (15)
Functional Reach <15cm (higher score = lower risk of falls)	25 (7)	30 (7)	42 (9)	97 (8)
Falls Efficacy Scale International score $\geq 11$ (higher score = more concerned about falling)	66 (18)	61 (15)	82 (18)	209 (17)

#### 4.5 Functional assessment outcomes at baseline

When compared with normative data from older, healthy populations, ProAct65+ baseline functional assessments revealed functional levels lower than published averages for all assessments, despite the significantly higher percentage meeting the UK guidelines on physical activity than the general UK population (Table 4.3). When baseline psychological assessments Short FES-I and CONFbal were compared with published normative values, the ProAct65+ population appeared less fearful of falling and more confident in maintaining their balance (Table 4.3). However, the comparator mean value for CONFbal was derived from a population attending day centres, so the greater confidence of ProAct65+ participants was anticipated. The full range of baseline assessment data compared with normative scores can be viewed in the ProAct65+ report (Iliffe et al., 2014).

**Table 4.3,** Baseline falls-related measure means compared to normative data

Outcome measure	ProAct65+ Mean (SD) [95% CI]	Normative mean (SD) [95% CI]	Normative reference
Timed Up and Go (seconds)	11.08 (5.94) [10.74, 11.42]	9.4 [8.9, 9.9]	Bohannon, 2006
30 second Chair Rise (number of stands)	Women 10.03 (3.02) [10.15, 10.49] Men 11.06 (3.54) [10.86, 11.26]	Women 12.7 (4.0) [12.45, 12.95] Men 14.2 (4.6) [13.91, 14.49]	Jones et al., 1999
Functional Reach (cm)	Men 26.34 (8.38) [25.86, 26.82] Women 24.93 (7.77) [24.48, 25.38]	Men 33.43 (1.55) [33.04, 33.82] Women 26.59 (3.53) [25.71, 27.47]	Duncan et al., 1990
Romberg (scored out of 28, high score is good)	20.19 (6.98)	None published as a score	
CONFbal (scored between 10 and 30, low score is good)	12.63 (3.98) [12.40, 12.86]	17.59 (7) [15.50, 19.68]	Simpson et al., 2009 (population attending day centres)
FES-I (range for each item = 1-4, 1=not at all concerned, 4=very concerned)  Items are matched by question although item numbers differed	Item 1 1.18 (0.54) Item 2 1.37 (0.72) Item 3 1.14 (0.49) Item 4 1.44 (0.76) Item 5 1.41 (0.75) Item 6 1.34 (0.69) Item 7 1.18 (0.54) Mean = 1.29 (0.64) [1.25, 1.33]	Item 2 1.50 (0.81) Item 4 2.09 (1.09) Item 6 1.49 (0.79) Item 7 2.06 (1.08) Item 9 2.14 (1.11) Item 15 2.46 (1.16) Item 16 1.85 (1.06) Mean = 1.94 (1.01) [1.86, 2.02]	Yardley et al., 2005

#### 4.6 ProAct65+ continuation of exercise

As described in the full ProAct65+ trial report, there was a significant increase of minutes of moderate to vigorous intensity physical activity (MVPA) per week at 12 months post intervention in the FaME group compared to usual care (UC) and a small non-significant increase in the OEP group compared to usual care. Between baseline and 12 months post-intervention the change in proportions meeting or exceeding the

MVPA requirement for health (150 minutes) by group was as follows; 40% to 49% in the FaME arm, 41% to 43% in the OEP arm and 37.5% to 38.0% in the usual care arm. The proportion reporting 150 minutes or more of MVPA per week at 12 months following the intervention in the FaME group was statistically significantly higher than in the usual care group (adjusted odds ratio (AOR) 1.78, 95% confidence interval (CI) 1.11 to 2.87;  $p=0.02$ ). The proportion reporting 150 minutes or more of MVPA per week at 12 months following the intervention in the OEP group was not statistically significantly different compared with the usual care group (AOR 1.17, 95% CI 0.72 to 1.92;  $p=0.52$ ) (Iliffe et al., 2014).

#### **4.7 Adherence**

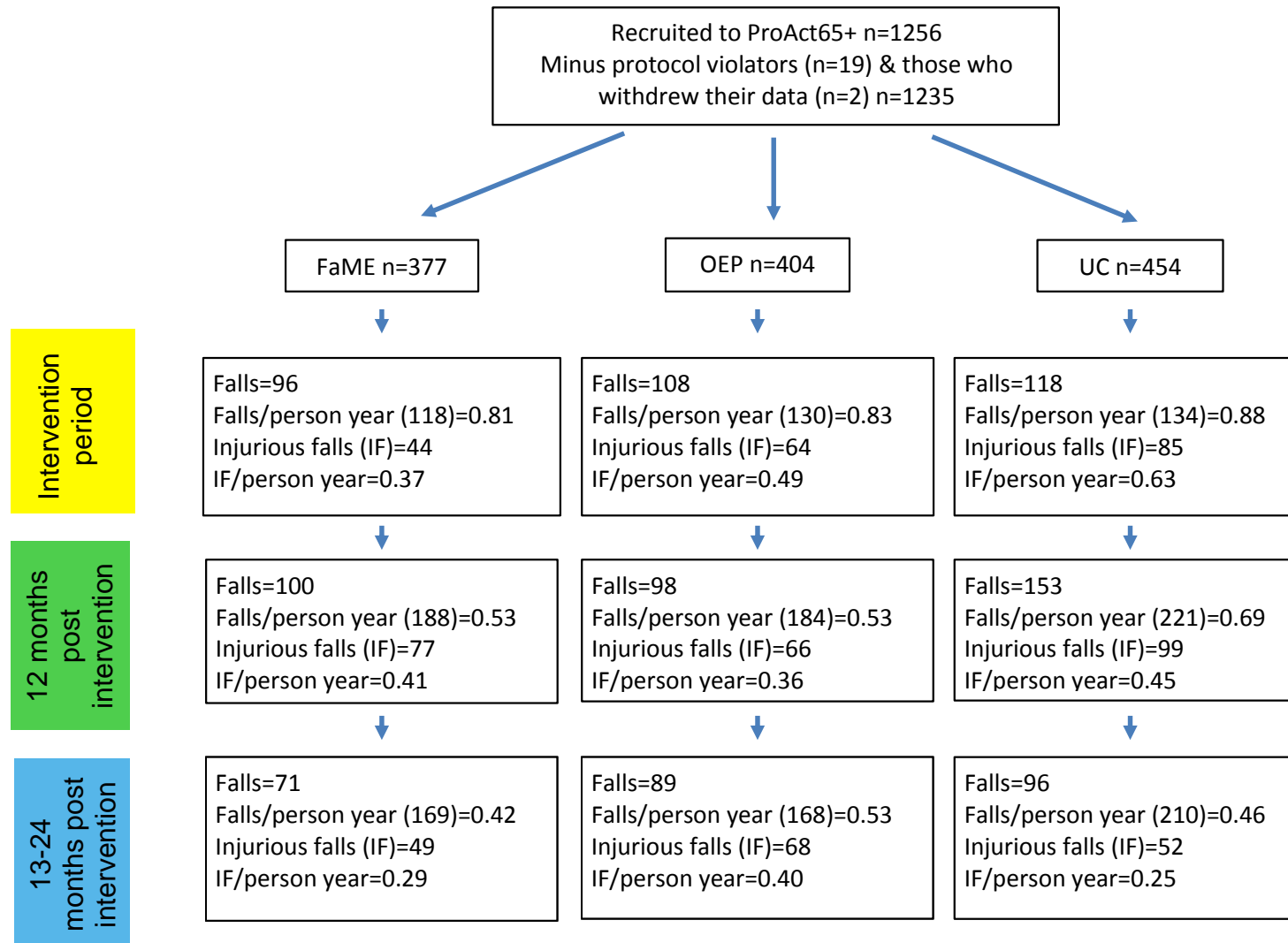
150 participants (40%) in the ProAct65+ FaME group attended 75% or more of classes. In the OEP group, 149 (37%) subjects reported that they achieved 75% or more of the home exercise prescription. This will be discussed further in Chapter 6, section 6.3.3.

#### **4.8 Falls; Intervention and Follow Up**

322 falls were reported during the 24-week intervention period, 351 in the first post-intervention year and 256 in the second year. The number of falls, and the number of falls that were injurious, by group for each time point are displayed in Figure 4.2, along with the corresponding number of person years. Person years take into account attrition and missing data, and also time at risk. Person time at risk was similar between groups at all time points.

The 322 falls during the intervention period were reported by 172 fallers; 50 (13%), 56 (14%) and 66 (15%) fallers were from the FaME, OEP and UC arms, respectively. The average number of falls per person during the intervention period was 0.25 in the FaME group, 0.27 in OEP and 0.26 in UC. There was no significant difference between either of the exercise interventions' falls incidence rate and that of UC during the intervention period (Table 4.4).

**Figure 4.2,** Falls flow chart: person years reflect attrition, missing data and time at risk



The 351 falls in the 12 months following the close of the interventions were reported by 194 fallers; 59 (16%), 59 (15%) and 76 (17%) fallers came from the FaME, OEP and UC arms, respectively. Average number of falls per person was 0.27 in the FaME group, 0.24 in OEP and 0.34 in UC. In this phase there was a 26% reduction in falls in the FaME group compared with UC (Table 4.4) and a statistically non-significant 24% reduction in the OEP arm (FaME: IRR=0.74, 95% CI 0.55, 0.99,  $p=0.04$ , OEP: IRR=0.76, 95% CI 0.53, 1.09,  $p=0.14$ ) (Table 4.4). We performed a post-hoc analysis to explore the poorer effect of the OEP intervention. When only those patients achieving 75% or more of the OEP intervention were compared with UC, there was a 59% reduction in falls during the intervention (IRR=0.41, 95%CI 0.24, 0.70:  $p=0.001$ ) and a 57% reduction in falls in the 12 months following the close of the intervention (IRR=0.43, 95%CI 0.21, 0.87:  $p=0.02$ ) (Table 4.4).

In the second year following the discontinuation of interventions, the effect of the interventions on falls rate was lost (FaME: IRR=0.94, 95% CI 0.62, 1.41,  $p=0.76$ , OEP: IRR=1.04, 95% CI 0.69, 1.55,  $p=0.86$ ). Given that there was a statistically significant reduction in falls during the year following the end of the FaME intervention, followed by a loss of this effect in the second year, a post-hoc supplementary analysis of this group was carried out to further investigate the second year. We found that when those in the FaME group who continued to report 150 minutes of MVPA per week into the second post-intervention year were compared to those in the FaME group who did not report maintaining their physical activity, there continued to be a significant reduction in falls incidence (IRR=0.49, 95% CI 0.30, 0.79;  $p=0.004$ ) (Table 4.4).

**Table 4.4, Falls Incident Rates & Rate Ratios**

	FaME	OEP	Usual care (UC)
<b>During the intervention</b>			
Number of fallers	50	56	66
Falls per person year*	0.81	0.83	0.88
Falls Rate Ratio (95% CI) (compared to UC)	0.91 (0.54, 1.52) p=0.72	0.93 (0.64, 1.37) p=0.72	Ref
Per protocol analysis (OEP only): Falls Rate Ratio (95% CI) (OEP 75% adherence compared to UC)	NA	0.41 (0.24, 0.70) p=0.001	Ref
Injurious Falls per person year*	0.37	0.49	0.63
Injurious Falls Rate Ratio (95% CI) (compared to UC)	0.55 (0.31,0.96) p=0.04	0.77 (0.50,1.20) p=0.25	Ref
<b>First year post intervention</b>			
Number of fallers	59	59	76
Falls per person year*	0.53	0.53	0.69
Falls Rate Ratio (95% CI) (compared to UC)	0.74 (0.55, 0.99) p=0.04	0.76 (0.53, 1.09) p=0.14	Ref
Injurious Falls per person year*	0.41	0.36	0.45
Injurious Falls Rate Ratio (95% CI) (compared to UC)	1.00 (0.70,1.45) p=0.98	0.69 (0.43,1.10) p=0.12	Ref
<b>Combined intervention and first year post-intervention period</b>			
Falls per person year	0.64	0.66	0.76
Falls Rate Ratio (95% CI) (compared to UC)	0.81(0.59,1.10) p=0.18	0.86 (0.62,1.19) p=0.36	Ref
Injurious Falls per person year	0.40	0.41	0.52
Injurious Falls Rate Ratio (95% CI) (compared to UC)	0.73 (0.54,0.99) p=0.05	0.74 (0.50,1.10) p=0.13	Ref
<b>Second year post intervention</b>			
Falls per person year*	0.42	0.53	0.46
Falls Rate Ratio (95% CI) (compared to UC)	0.94 (0.62,1.41) p=0.76	1.04 (0.69,1.55) p=0.86	Ref
Falls per person year (FaME only)	<150 mins MVPA 0.59 ≥150 mins MVPA 0.30	NA	NA

	FaME	OEP	Usual care (UC)
Within-group analysis (FaME only): Falls Rate Ratio (95% CI) (<150 mins MVPA compared to ≥150 mins MVPA)	0.49 (0.30, 0.79) p=0.004	NA	NA
Injurious Falls per person year*	0.29	0.40	0.25
Injurious Falls Rate Ratio (95% CI) (compared to UC)	1.44 (0.78,2.64) p=0.24	1.50 (0.89,2.53) p=0.13	Ref

\* Person years are displayed in Figure 4.2

### Injurious falls

Injurious falls during the intervention totalled 44 in the FaME group, 64 in the OEP group and 85 in UC, as reported by patients. Negative binomial modelling revealed significantly fewer injuries in the FaME group compared with UC during the 24-week intervention (Table 4.4) and in the combined intervention and first year post-intervention period (Table 4.4). In the second year following the close of interventions, the effect of FaME on injurious falls rate was lost (Table 4.4). The difference in number of injuries in the OEP group compared with UC was not significant at any time-point, though there was a statistically non-significant reduction (IRR 0.74, 95% CI 0.50, 1.10; p=0.13) in the combined intervention and first year post-intervention period. The injurious falls rate was lower in all groups in the second year post intervention than in the intervention period, with the greatest reduction being in the usual care group.

### 4.9 Falls diary data

Despite telephone call reminders from the researcher to return diaries, diary return was poor, resulting in missing falls data. Overall, 62% of intervention diaries were returned. 595 (48%) patients returned all 6 diaries, 345 (28%) did not return any. 35%, 37% and 41% of diaries were missing in the OEP, FaME and UC groups, respectively. There was no association between returning diaries and gender nor age, but fallers were less likely to return diaries than non-fallers (Perry et al., 2012). Those patients who returned all 6 diaries had a falls rate of 0.67 falls/person year, but those who

returned between one and three diaries had a rate of 1.59 falls/person year. I will discuss this further in Chapter 6, section 6.4.

#### **4.10 Falls risk factors**

Two risk factors for falls; balance and lower limb strength, were measured by the functional assessments (TUG, Functional Reach, Romberg and timed chair rise), at baseline and at the end of the intervention period. As reported elsewhere, there were no statistically significant changes in TUG, Functional Reach or timed chair rise following the intervention (Iliffe et al., 2014). The post-intervention Romberg data were not analysed, nor included in the full HTA report, due to some measurement errors for some patients. Fear of falling was measured by Short FES-I and CONFbal, at the end of the interventions and at all subsequent follow-up points. The Falls Risk Assessment Tool (FRAT) was also measured at the end of the interventions and at all subsequent follow-up points. As reported elsewhere, there were no statistically significant changes in any of these measures at 12 months post intervention, with the exception of CONFbal, which was significantly improved in both intervention arms compared with UC (Iliffe et al., 2014). There were no significant changes in any of these measures at 24 months post intervention. Table 4.5 shows functional assessment, fear of falling, and falls risk assessment scores at all time-points. Timed up and go (TUG) scores were transformed to log values as the data were not normally distributed.



**Table 4.5, Secondary falls outcomes by group (data from ProAct65+ trial report)**

Outcome			FaME	OEP	UC
Baseline	Log-TUG (seconds)	Mean SD	2.33 0.34	2.33 0.34	2.35 0.32
	Functional Reach (cm)	Mean SD	25.60 6.98	25.57 7.43	24.68 7.43
	Chair rise (total number of stands)	Mean SD	10.48 3.64	10.27 2.81	10.40 3.31
	Short FES-I (score)	Mean SD	8.99 3.56	8.89 3.49	9.36 4.08
	CONFbal (score)	Mean SD	12.63 3.98	12.48 3.76	12.55 3.93
	FRAT (score)	Mean SD	0.89 0.90	0.98 0.90	1.03 0.96
12 months post intervention	Log-TUG (seconds)	Mean SD	2.25 0.30	2.27 0.27	2.28 0.27
		Estimate* 95% CI P-value	-0.008 -0.064,0.048 0.775	-0.011 -0.066,0.044 0.700	Ref
	Functional Reach (cm)	Mean SD	26.99 7.28	26.84 7.64	27.13 6.82
		Estimate 95% CI P-value	-0.644 -2.583,1.295 0.515	-1.055 -3.031,0.921 0.295	Ref
	Chair rise (total number of stands)	Mean SD	11.62 3.77	11.40 3.35	11.86 3.57
		Estimate 95% CI P-value	-0.644 -2.583,1.295 0.515	-1.055 -3.031,0.921 0.295	Ref
	Short FES-I (score)	Mean SD	9.20 4.56	9.09 4.19	8.94 3.66
		Estimate 95% CI P-value	0.102 -0.653,0.856 0.792	0.045 -0.740,0.831 0.910	Ref
	CONFbal (score)	Mean SD	12.13 3.65	12.23 3.71	12.38 4.05
		Estimate 95% CI P-value	-0.529 -0.998, -0.061 0.027	-0.545 -1.033, -0.057 0.029	Ref
	FRAT (score)	Mean SD	0.93 0.94	1.00 0.95	0.99 0.91

Outcome			FaME	OEP	UC
		Estimate 95% CI P-value	-0.004 -0.160,0.152 0.960	0.030 -0.127,0.189 0.708	Ref
13-24 months post intervention	Short FES-I (score)	Mean SD	9.08 4.17	8.78 3.54	9.37 4.23
		Estimate 95% CI P-value	-0.46 -1.22,0.29 0.230	-0.70 -1.48,0.08 0.077	Ref
	CONFbal (score)	Mean SD	12.45 4.05	12.13 3.33	12.55 4.02
		Estimate 95% CI P-value	-0.40 -0.96,0.16 0.160	-0.54 -1.10,0.03 0.065	Ref

\*Estimate=estimate of difference in mean between intervention and usual care from multiple linear regression model

#### 4.11 Chapter conclusions

This chapter has described the ProAct65+ population demographics and falls characteristics. As my therapist and dose effects (TDE) group was derived from the FaME group of the ProAct65+ population this chapter has described the larger population in which the TDE group is nested. A comparison between the TDE group, and those in the FaME arm who were excluded from the TDE group, can be found in the next chapter; Results 2; Dose and therapist effects. This chapter has also presented the ProAct65+ falls analysis findings that underpin my TDE study. As my work investigates the effect of the therapist and dose on patient falls outcomes, it was important to describe the falls outcomes in the larger ProAct65+ group first. As stated earlier, the falls analysis intellectual property is not mine, but I led on the falls analysis for the ProAct65+ team.

## **Chapter 5**

### **Results 2; Dose and therapist effects**

#### **5.1 Chapter summary**

In this chapter I describe all the findings relating to the therapist and dose effects analysis, starting with descriptive statistics to characterise the therapists themselves, and the patients seen by the therapists; a sub-set of the ProAct65+ FaME group. Following this, the bivariate analyses are presented, and finally the multilevel modelling analyses. The bivariate analyses aimed to describe any relationships between patient outcomes (falls and functional measures of falls risk factors) and therapist characteristics or dose. The multilevel modelling confirmed any therapist or dose effects shown in the simpler analyses, with the advantage of being designed for clustered data, thus ensuring that any effects were not over-estimated. The multilevel modelling also investigated whether specific therapist characteristics were individually associated with specific patient outcomes, and the magnitude of the effect.

#### **5.2 Therapist characterisation**

The 12 PSIs who were deployed in the main trial, and were therefore those included in the therapist effect (TE) analysis, are described in Table 5.1. Six worked at each trial site. The total number of patients they taught within trial FaME group classes varied from 8 to 71 patients, depending on the PSIs' availability and therefore the number of classes they were asked to deliver. The average exercise class size was 6 patients, with a range of 3 to 10 patients.

**Table 5.1, PSI characteristics**

Characteristic	Outcome
Age	Mean=42, Range=33-60 years
Gender	2 men, 10 women
FaME delivery experience prior to the trial?	6 (50%) had previous experience
PSI qualification attained before the trial?	8 (67%) attained before the trial
Professional background	2 physiotherapists, 10 exercise professionals
Attendance (% of intervention they taught themselves)	Range=64-100%
Quality Assurance 'rating'*	1 x low, 5 x moderate, 6 x high

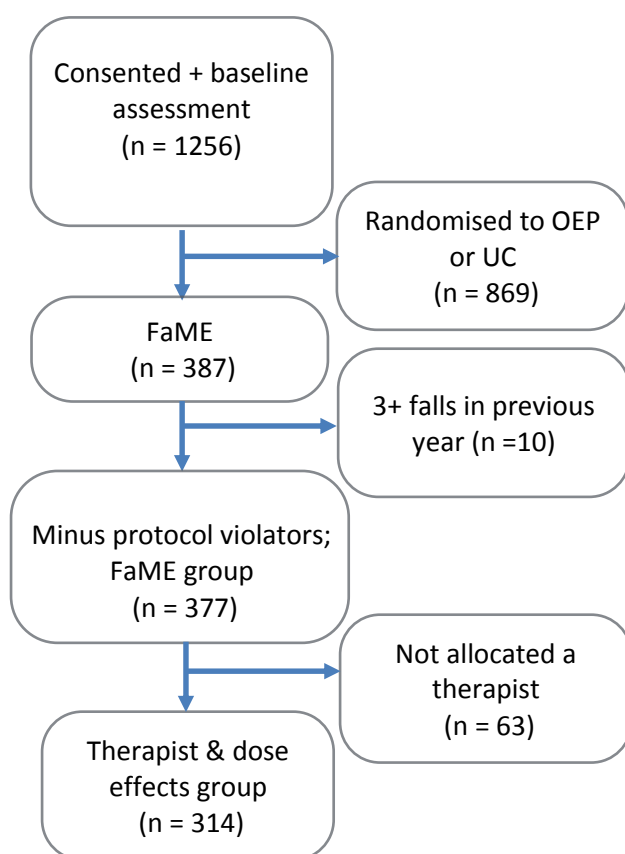
\* See Chapter 3, section 3.4.5.3 for definition & calculation method

### 5.3 Baseline characteristics of therapist effect study population

Therapist (and dose) effects were analysed in the FaME group only as this was the only group who received regular (weekly) exercise instruction from trained specialist exercise professionals. Of the 387 patients who were assigned to the FaME arm, 314 (81%) were allocated a class, and therefore, a PSI. The remainder either 1) were protocol violators (more than 2 falls in the year preceding the study) and therefore excluded from the falls and TE analyses (n=10), 2) withdrew on allocation (n=31), 3) were un-contactable (neither answered the telephone nor responded to messages asking them to contact the research team) following baseline assessment and randomisation (n=2) or 4) excluded by their GP (n=30). As these individuals had no contact with a therapist and therefore could not be assigned to a PSI level group in the multilevel model analysis, they were excluded from the therapist and dose effects (TDE) analysis (Figure 5.1). The 314 patients included in the TDE analysis had a mean age of 73 (range 65-92), a median of 2 co-morbidities and a median number of 4 medications. 198 (63%) were female. 67 (21%) reported 1 or 2 falls in the year preceding the study. The total number of falls reported in the year preceding the study was 82. 298 (95%) reported being able to use public transport easily and 50 (16%) reported using a walking aid. 137 (44%) stated they already exercised regularly each week and 117 (37%) were achieving 150 minutes or more of moderate (or higher)

intensity physical activity per week i.e. they were meeting or exceeding the Department of Health guidelines regarding sufficient physical activity to benefit health. In terms of falls risk-related measures, the 314 patients presented as in Table 5.2.

**Figure 5.1,** CONSORT diagram showing flow of patients through trial



**Table 5.2,** Therapist and dose effects (TDE) group baseline falls & falls risk factors compared to those excluded from the TDE group

Outcome measure	TDE group (n=314), n (%)	Non-TDE group (n=63), n (%)
Reported (1 or 2) falls in previous year	67 (21%)	15 (24%)
Average number of falls per person in previous year	0.26	0.27
Falls Risk Assessment Tool $\geq 3$	14 (4%)	2 (3%)
Timed Up and Go > 13.5 seconds	37 (12%)	14 (22%)
Functional Reach < 15cm	11 (4%)	12 (19%)
Confidence in Maintaining Balance score $\geq 20$	16 (5%)	7 (11%)
Falls Efficacy Scale International score $\geq 11$	47 (15%)	14 (22%)

The TDE group, compared to those excluded from the TDE group (non-TDE group), scored better in all functional assessments and measures of falls risk factors, such as leg power and balance. The baseline falls rate, however, was very similar in both groups; 0.26 and 0.27 in the TDE group and the non-TDE group, respectively (Table 5.2).

#### 5.4 Participant evaluation

A questionnaire designed for the trial was issued to participants in the intervention group to gather their evaluation of the exercise programmes. 196/377 (52%) of the patients in the FaME group returned their exercise programme evaluation questionnaire. In the TDE group (n=314), patient evaluation data were available for 62% (n=196/314) of the group. In other words, all 196 patients who returned the questionnaire are included in the TDE group. Table 5.3 compares the characteristics of those in the FaME group who did and those who did not return the questionnaire. Those who did not return the questionnaire were on average less functionally able, had poorer balance and leg power, fell more and were less likely to have attended the intervention than those who did return the questionnaire. The missing data for CONFbal and Short FES-I (Table 5.3) were due to 45/377 (12%) patients in the FaME

group not returning their baseline questionnaire booklet. The missing data for TUG (Table 5.3) were due to insufficient space in the GP practice to administer the TUG assessment. The missing data for falls (Table 5.3) were due to patients not returning their self-completion falls diaries.

**Table 5.3,** Comparison of those in the FaME group who did/did not return the evaluation questionnaire

	<b>Patients who returned the evaluation questionnaire n=196</b>	<b>Patients who did not return the evaluation questionnaire n=182</b>
<b>Mean baseline TUG in seconds (SD)</b>	9.80 (3.54) n=171	11.96 (5.80) n=157
<b>Mean baseline functional reach in cm (SD)</b>	26.76 (6.96) n=187	24.67 (6.74) n=174
<b>Mean baseline CONFbal score (SD)</b>	11.78 (3.05) n=180	13.46 (4.50) n=142
<b>Mean baseline Short FES-I score (SD)</b>	8.59 (3.55) n=181	9.29 (3.32) n=143
<b>Mean percentage class attendance (SD)</b>	74 (22) n=195	30 (34) n=179
<b>Mean falls rate (falls/person year) during intervention (SD)</b>	0.67 (2.13) n=193	1.28 (3.90) n=84

Of those who returned the questionnaire, 98% stated that they attended the exercise classes, 96% found the classes very enjoyable or somewhat enjoyable, 74% felt the intensity of classes was 'about right', 100% had confidence in their class instructor and 86% stated that they did their home exercise programme. Selected information from the questionnaires was used in the therapist effect (TE) analysis. Only responses from two participant evaluation questions were used in the TE analysis: 1) 'How did you find the (intensity of the) exercise sessions?' and 2) 'Were the sessions enjoyable?' as these questions were considered to provide data relating to the skill of the therapist. The responses are shown in Table 5.4. It was also planned to use the question 'Did you have confidence in your instructor?' but all patients who answered the questionnaire responded positively, so there was no difference between therapists to investigate.



**Table 5.4,** TDE group evaluation questionnaire responses used in TE analyses

Question	Response	n (%)
How did you find the (intensity of the) exercise sessions? n=196	Far too easy	11 (6)
	A little too easy	36 (18)
	About right	143 (73)
	A little too hard	6 (3)
	Far too hard	0 (0)
Were the sessions enjoyable? n=195	Very enjoyable	150 (77)
	Somewhat enjoyable	38 (19)
	Neither enjoyable nor not enjoyable	5 (3)
	Not really enjoyable	2 (1)
	Not enjoyable at all	0 (0)

## 5.5 Sample size

I calculated that a final sample size of 4,192 participants would have been necessary to detect a difference in the main falls outcome in my therapist effect study. ProAct65+ reported a 26% reduction in falls in the FaME group in the first post-intervention year. We also found that 24% of the recruited population had fallen in the 12 months preceding the study. A bespoke therapist effects study would therefore be designed to detect a 26% reduction (from 24% to 18%) in the proportion of participants who have at least one fall during follow-up. A sample size of 690 per group ( $\alpha = 0.05$ , power = 80%) would be necessary for this. ('Group', in the context of a therapist effect study sample size calculation, refers to patients exercised by more skilled therapists and those exercised by less skilled therapists, although, in reality, therapist skill is on a continuum.) The sample must then be further inflated to allow for the clustering effect of patients within therapists. There were 314 patients in my therapist effect group, exercised by 12 PSIs. On average, therefore, each PSI exercised 26 patients. The ICC (the degree to which patients clustered by PSI resemble each other in terms of the

primary falls outcome) from my therapist effect study results was 0.05 (see section 5.6.4.3). In a bespoke therapist effect study the design effect, therefore, would be:

$$1 + (m-1) \times ICC = 1 + 25 \times 0.05 = 2.25$$

Inflating the sample size ( $2 \times 690 = 1380$ ) by a factor of 2.25 gives a sample size of 3,105 patients. With an estimated dropout rate of 35%, a final sample size of 4,192 would be needed. This large sample is needed to detect the very limited impact of different therapists on falls. My actual sample size, as indicated earlier, was 314, and therefore my study was underpowered.

## **5.6 Therapist and dose effects analysis results**

### **5.6.1 Data cleaning**

Histograms were plotted for the outcome variables and can be found in the Appendix 4. The data were not all normally distributed. For example, TUG data were right-skewed, Romberg balance data were left-skewed, and the falls data demonstrate poisson distribution (which is 'usual' for falls data). Data that are not normally distributed can be transformed, for example, log transformation of right-skewed data. However, the central limit theorem suggests that inferences can be reasonable for large sample sizes even if data do not follow the classic Gaussian distribution (Lumley et al., 2002). As well as this, the t-tests, ANOVAs and correlations are superseded by the multilevel modelling which allows for clustering, so the bivariate analyses results would be interpreted with caution, even if outcome data all followed normal distributions. For these reasons, I decided not to transform the data nor to conduct sensitivity analyses to check that results were the same using transformed data compared to using untransformed data.

### **5.6.2 Bivariate analyses**

The therapist and dose effects analysis explores possible associations between falls rate or falls risk factors and therapist characteristics or dose of exercise. To verify the data, mean and standard deviations, as well as 95% confidence intervals, for all

dependent variables by therapist characteristic were undertaken, and are shown in Tables 5.5 to 5.8.

**Table 5.5,** Mean falls rate, standard deviation & confidence intervals, by PSI characteristic

Therapist characteristic		Mean falls rate in falls per person year during intervention (SD) [95% CI]	Mean falls rate in falls per person year post-intervention (SD) [95% CI]
PSI gender	Male	0.60 (1.10) n=20 [0.13, 1.15]	0.14 (0.36) n=14 [0.00, 0.36]
	Female	0.85 (2.82) n=252 [0.53, 1.20]	0.51 (1.10) n=214 [0.37, 0.66]
PSI professional background	Exercise instructor	0.83 (2.74) n=242 [0.49, 1.17]	0.47 (1.02) n=200 [0.33, 0.62]
	Physiotherapist	0.78 (2.74) n=30 [0.08, 1.88]	0.63 (1.39) n=28 [0.19, 1.21]
PSI experience	Novice	0.72 (2.66) n=192 [0.38, 1.82]	0.53 (1.09) n=165 [0.38, 0.71]
	Experienced	1.10 (2.90) n=80 [0.53, 1.82]	0.39 (1.03) n=63 [0.16, 0.68]
PSI training	Before trial	0.80 (2.52) n=180 [0.45, 1.20]	0.44 (0.98) n=148 [0.30, 0.61]
	During trial	0.89 (3.12) n=92 [0.32, 1.60]	0.57 (1.23) n=80 [0.32, 0.86]
PSI 'quality'	Low	0.92 (3.19) n=27 [0.07, 2.35]	0.67 (1.35) n=21 [0.22, 1.35]
	Moderate	1.21 (3.40) n=91 [0.59, 1.98]	0.37 (0.77) n=77 [0.21, 0.56]
	High	0.58 (2.14) n=154 [0.28, 0.94]	0.53 (1.17) n=130 [0.34, 0.73]

Table 5.5 shows mean falls rates in falls per person year by PSI characteristic. Falls rates are separated into two time-points; during the intervention and in the 12 months post-intervention. The mean falls rate after the intervention appears to be lower for those patients taught by a male PSI. Otherwise, the confidence intervals associated

with mean falls rates for therapist characteristic categories overlap, suggesting that there are no differences in falls rates according to the characteristics of the therapist.

**Table 5.6,** Mean falls rate, standard deviation and confidence intervals, by patient evaluation of exercise

Evaluation type		Mean falls rate in falls per person year during intervention (SD) [95% CI]	Mean falls rate in falls per person year post-intervention (SD) [95% CI]
Patient enjoyment of exercise	very enjoyable	0.61 (2.05) n=149 [0.32, 0.98]	0.39 (0.94) n=147 [0.24, 0.54]
	somewhat enjoyable	0.46 (1.02) n=37 [0.16, 0.78]	0.59 (1.09) n=31 [0.26, 1.00]
	neither enjoyable nor not enjoyable	4.00 (6.16) n=5 [0.00, 11.26]	0.93 (1.30) n=5 [0.00, 2.33]
	not really enjoyable	2.00 (0.00) n=2 [2.00, 2.00]	4.00 n=1 [4.00, 4.00]
	Not enjoyable at all	n=0	n=0
Patient-rated intensity of exercise	far too easy	0.36 (0.81) n=11 [0.00, 0.89]	0.70 (1.34) n=10 [0.00, 1.67]
	a little too easy	0.81 (2.61) n=36 [0.91, 1.78]	0.50 (0.93) n=34 [0.21, 0.83]
	about right	0.68 (2.10) n=141 [0.39, 1.09]	0.43 (1.02) n=135 [0.26, 0.60]
	a little too hard	0.40 (0.99) n=6 [0.00, 1.44]	0.72 (0.85) n=6 [0.00, 1.47]
	far too hard	n=0	n=0

Table 5.6 shows mean falls rates by patients' evaluation of the exercise intervention with regard to the intensity and their enjoyment of the sessions. The confidence intervals associated with mean falls rates for intensity and enjoyment categories overlap, suggesting that there are no differences in falls rates in groups of patients according to their evaluation of the intervention.

Table 5.7a shows mean baseline functional assessment scores (TUG, functional reach, chair rise & Romberg balance) by PSI characteristic. Despite randomisation being

clustered by general practice, there appears to be little difference in baseline score between the groups for most functional assessments.

**Table 5.7a**, Baseline functional assessment score mean, standard deviation and confidence intervals, by PSI characteristic

Therapist characteristic		Mean baseline TUG in seconds (SD) [95% CI]	Mean baseline functional reach in cm (SD) [95% CI]	Mean baseline Romberg score (SD) [95% CI]	Mean baseline number of chair stands (SD) [95% CI]
PSI gender	Male	11.33 (11.99) n=16 [7.77, 18.75]	30.75 (6.06) n=20 [28.11, 33.44]	21.45 (7.49) n=22 [18.12, 24.30]	10.77 (4.14) n=22 [9.11, 12.62]
	Female	10.60 (4.24) n=237 [10.12, 11.17]	26.14 (6.66) n=282 [25.34, 26.89]	21.01 (6.35) n=292 [20.27, 21.71]	10.65 (3.52) n=286 [10.19, 11.08]
PSI professional background	Exercise instructor	10.54 (5.02) n=221 [9.91, 11.27]	26.54 (6.74) n=270 [25.73, 27.33]	21.37 (6.03) n=282 [20.68, 22.02]	10.62 (3.48) n=277 [10.20, 11.06]
	Physiotherapist	11.42 (5.21) n=32 [9.76, 13.37]	25.66 (6.49) n=32 [23.41, 27.69]	18.13 (8.80) n=32 [15.06, 20.95]	11.00 (4.31) n=31 [9.61, 12.64]
PSI experience	Novice	10.84 (4.37) n=207 [10.31, 11.44]	26.14 (6.80) n=222 [25.26, 27.06]	20.68 (6.63) n=224 [19.79, 21.49]	10.56 (3.66) n=218 [10.07, 11.03]
	Experienced	9.77 (7.35) n=46 [8.27, 12.39]	27.28 (6.40) n=80 [25.79, 28.64]	21.93 (5.81) n=90 [20.63, 23.08]	10.89 (3.34) n=90 [10.22, 11.62]
PSI training	Before trial	10.14 (4.69) n=157 [9.45, 10.91]	26.80 (6.45) n=204 [25.89, 27.67]	21.24 (5.72) n=214 [20.41, 21.99]	10.32 (3.06) n=210 [9.90, 10.75]
	During trial	11.47 (5.51) n=96 [10.47, 12.53]	25.68 (7.20) n=98 [24.34, 26.99]	20.60 (7.72) n=100 [19.02, 22.10]	11.37 (4.38) n=98 [10.51, 12.31]
PSI 'quality'	Low	12.00 (8.35) n=25 [9.16, 15.67]	25.16 (5.69) n=29 [22.85, 27.28]	21.03 (5.85) n=29 [18.78, 22.92]	12.00 (5.03) n=29 [10.11, 13.83]
	Moderate	10.01 (3.20) n=88 [9.38, 10.74]	27.36 (7.16) n=107 [25.95, 28.70]	22.28 (4.84) n=116 [21.34, 23.20]	10.90 (3.75) n=115 [10.19, 11.64]
	High	10.81 (5.21) n=140 [10.05, 11.74]	26.07 (6.53) n=166 [25.09, 27.04]	20.18 (7.30) n=169 [19.04, 21.24]	10.25 (3.03) n=164 [9.79, 10.73]

Table 5.7b shows mean functional assessment scores (TUG, functional reach, chair rise & Romberg balance) by PSI characteristic from the immediate post-intervention meeting with the patient. Most of the confidence intervals associated with mean scores for therapist characteristic categories (for example, mean Timed-Up-and-Go for patients taught by male PSIs compared with those taught by female PSIs) overlap, suggesting that there are no differences in post-intervention functional assessment scores according to the characteristics of the therapist.

**Table 5.7b**, Post-intervention functional assessment score mean, standard deviation and confidence intervals, by PSI characteristic

Therapist characteristic		Mean post-intervention TUG in seconds (SD) [95% CI]	Mean post-intervention functional reach in cm (SD) [95% CI]	Mean post-intervention Romberg score (SD) [95% CI]	Mean post-intervention number of chair stands (SD) [95% CI]
PSI gender	Male	12.06 (10.31) n=8 [6.88, 22.26]	24.47 (5.71) n=17 [21.53, 27.16]	21.39 (5.52) n=18 [18.45, 23.90]	10.67 (3.99) n=18 [8.85, 12.47]
	Female	9.85 (3.27) n=190 [9.39, 10.32]	27.09 (7.34) n=226 [26.15, 28.09]	21.73 (4.92) n=230 [21.09, 22.34]	11.68 (3.76) n=228 [11.21, 12.19]
PSI professional background	Exercise instructor	9.85 (3.72) n=168 [9.34, 10.46]	26.98 (7.46) n=213 [25.99, 28.04]	21.54 (4.90) n=219 [20.94, 22.16]	11.54 (3.77) n=217 [11.05, 12.07]
	Physiotherapist	10.43 (4.12) n=30 [9.11, 11.99]	26.42 (5.70) n=30 [24.20, 28.40]	22.97 (5.33) n=29 [20.87, 24.75]	12.10 (3.89) n=29 [10.73, 13.61]
PSI experience	Novice	10.03 (3.41) n=163 [9.52, 10.57]	27.04 (7.52) n=178 [25.95, 28.13]	21.50 (4.93) n=177 [20.77, 22.20]	11.88 (4.01) n=176 [11.34, 12.47]
	Experienced	9.51 (5.19) n=35 [8.13, 11.49]	26.55 (6.51) n=65 [24.96, 28.22]	22.23 (5.02) n=71 [21.06, 22.28]	10.91 (3.02) n=70 [10.19, 11.64]
PSI training	Before trial	9.89 (3.84) n=111 [9.25, 10.69]	26.82 (7.57) n=156 [25.67, 28.00]	21.50 (4.92) n=162 [20.82, 22.28]	10.99 (3.41) n=160 [10.42, 11.57]
	During trial	9.99 (3.71) n=87 [9.24, 10.77]	27.06 (6.90) n=87 [25.60, 28.62]	22.10 (5.03) n=86 [20.97, 23.11]	12.74 (4.16) n=86 [11.89, 13.62]
PSI 'quality'	Low	10.17 (4.30) n=25 [8.62, 12.01]	28.16 (6.87) n=25 [25.59, 30.80]	22.40 (5.19) n=25 [20.38, 24.41]	13.04 (5.24) n=25 [11.11, 15.04]
	Moderate	9.05 (2.05) n=50 [8.47, 9.61]	26.17 (6.40) n=78 [24.87, 27.58]	21.87 (4.66) n=83 [20.84, 22.84]	11.18 (3.81) n=82 [10.37, 12.00]
	High	10.25 (4.15) n=123 [9.62, 11.05]	27.10 (7.76) n=140 [25.80, 28.47]	21.49 (5.10) n=140 [20.63, 22.30]	11.60 (3.40) n=139 [11.03, 12.11]

Table 5.8 shows mean post-intervention functional assessment scores by patient evaluation of exercise. Functional assessment mean scores are poorer in those who found the sessions too challenging. This observation does not give any indication of



whether patients' perceived intensity of the exercise is grouped by therapist. Numbers of patients in some groups are very low, for example, in the 'not really enjoyable' group.

**Table 5.8,** Post-intervention functional assessment score mean, standard deviation and confidence intervals, by patient evaluation of exercise

Evaluation type		Mean post-intervention TUG in seconds (SD) [95% CI]	Mean post-intervention functional reach in cm (SD) [95% CI]	Mean post-intervention Romberg score (SD) [95% CI]	Mean post-intervention number of chair stands (SD) [95% CI]
Patient evaluation; enjoyment	very enjoyable	9.62 (2.78) n=117 [9.14, 10.16]	26.96 (7.26) n=142 [25.86, 28.14]	21.80 (5.07) n=143 [20.92, 22.64]	11.84 (3.60) n=143 [11.25, 12.43]
	somewhat enjoyable	9.83 (4.05) n=32 [8.60, 11.35]	28.16 (7.11) n=37 [25.74, 30.39]	23.00 (4.59) n=36 [21.46, 24.41]	11.64 (4.19) n=36 [10.34, 13.12]
	neither enjoyable nor not enjoyable	8.50 (2.12) n=2 [7.00, 10.00]	17.50 (2.12) n=2 [16.00, 19.00]	20.50 (5.20) n=4 [14.00, 26.00]	13.50 (2.38) n=4 [11.00, 16.00]
	not really enjoyable	10.35 n=1	43.00 n=1	20.00 (7.07) n=2 [15.00, 25.00]	14.00 n=1
Patient evaluation; intensity	far too easy	7.89 (1.87) n=10 [6.72, 9.07]	35.90 (4.04) n=10 [33.60, 38.67]	25.20 (2.62) n=10 [23.56, 26.86]	15.90 (4.93) n=10 [13.40, 19.83]
	a little too easy	8.29 (1.61) n=28 [7.71, 8.93]	27.03 (7.23) n=33 [24.56, 29.52]	23.57 (4.00) n=35 [22.25, 24.82]	13.29 (3.66) n=35 [12.17, 14.48]
	about right	10.09 (3.32) n=110 [9.50, 10.74]	26.71 (7.25) n=134 [25.46, 28.00]	21.51 (4.95) n=135 [20.62, 22.29]	11.34 (3.26) n=134 [10.76, 11.89]
	a little too hard	13.39 (3.43) n=5 [10.18, 15.95]	21.83 (5.64) n=6 [16.83, 26.00]	17.00 (8.46) n=6 [10.00, 24.00]	7.33 (2.16) n=6 [5.67, 9.50]

Bivariate analyses were conducted for each predictor variable with each dependent variable separately to explore potential relationships between the paired variables. Correlation, t-test or Analysis of Variance (ANOVA) was selected depending on the

type of variables within the pair (categorical variable versus continuous variable, for example). Results are presented next, according to analysis type.

#### **5.6.2.1 Correlation analyses**

Correlation analysis was used to explore associations between pairs of continuous variables. Exercise dose, and two therapist characteristics; PSI age and PSI attendance (the percentage of the intervention they taught themselves), are explored in this order, looking at the effect on falls rate, followed by the effect on the post-intervention functional assessment outcomes.

##### **5.6.2.1.1 Exercise dose**

Figure 5.2 shows correlation between the patients' dose of the intervention (their total attendance) and falls rate at three time-points; during the intervention, during the 12 months following the intervention, and during the combined intervention and post-intervention period. All three scatter plots show a weak negative association; falls incidence decreases with increased dose of exercise. However, there are many patients without any falls irrespective of exercise dose, therefore, falls rate was also dichotomised into the group of patients who did not fall and the group who did experience one or more falls. The associated histograms are shown in Figure 5.3.

During the intervention, falling does not appear to be associated with exercise dose. In the 12 months following the intervention, and in the combined intervention and post-intervention period, however, those patients who achieved a higher dose of the intervention appear to fall slightly less. The proportion of patients falling increases with the length of the observation period.

Figure 5.4 shows correlation between the patients' dose of the intervention and post-intervention the functional assessment outcomes. Post-intervention chair rise, functional reach and Romberg balance scores do not appear to be affected by exercise dose. Timed-up-and-go has a weak negative association with dose; post-intervention TUG time decreases slightly with a higher dose of the intervention.

**Figure 5.2,** Correlation analyses for exercise dose (patient attendance) and falls rate

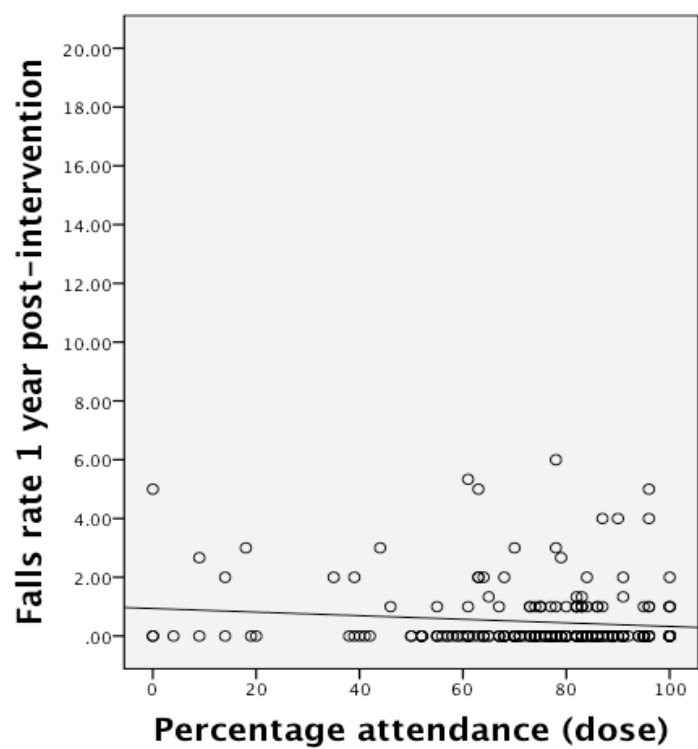
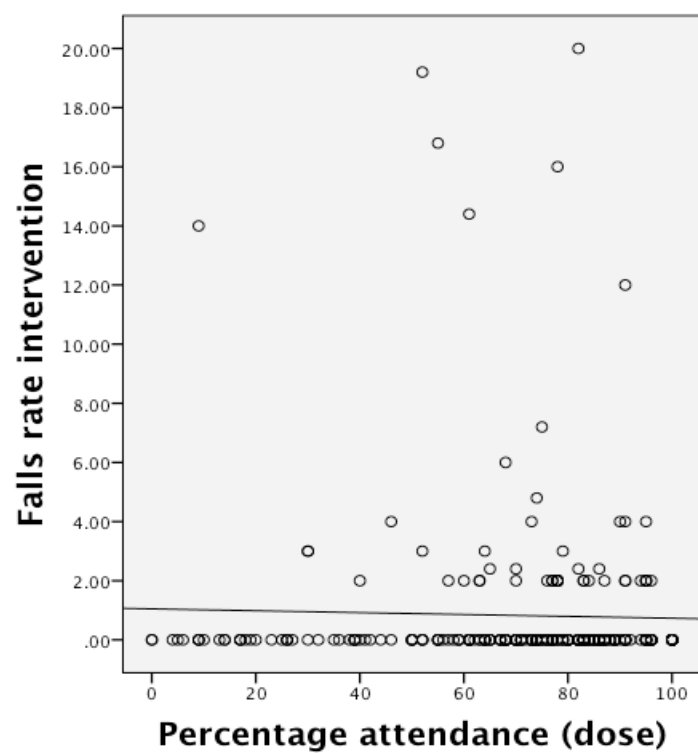
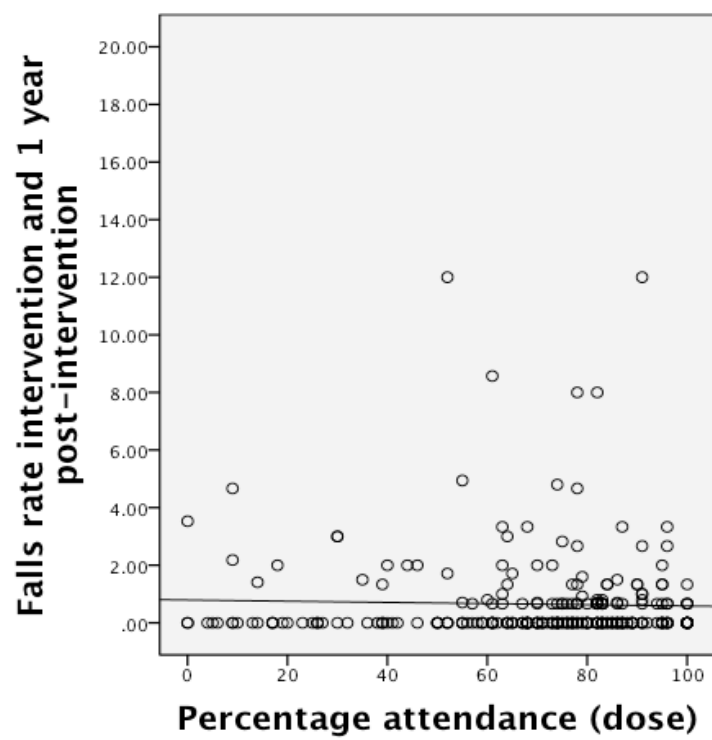
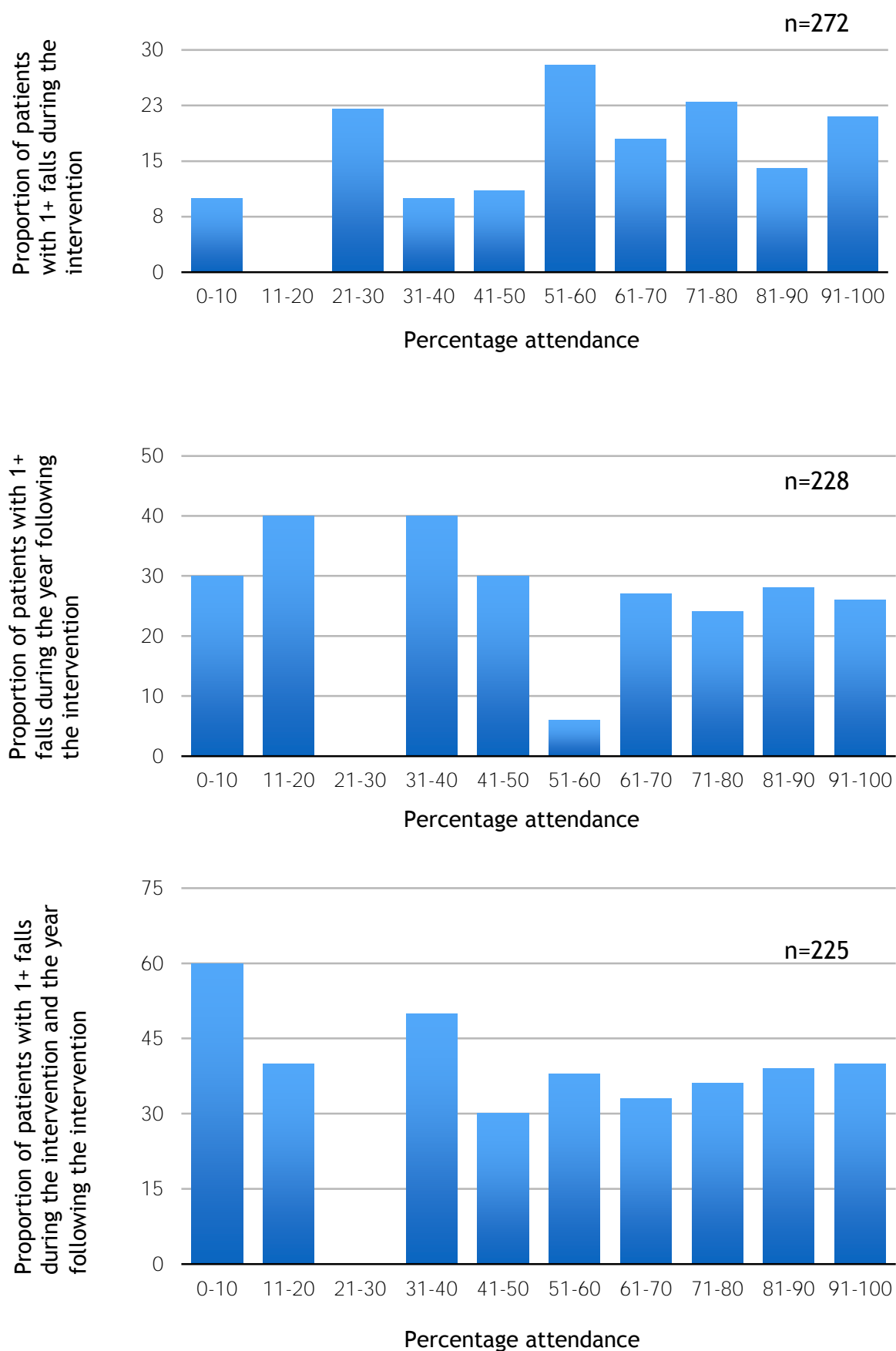


Figure 5.2, continued



**Figure 5.3, Proportion of fallers (1+ falls) by dose of exercise and time-point**



**Figure 5.4,** Correlation analyses for exercise dose (attendance) and functional outcomes

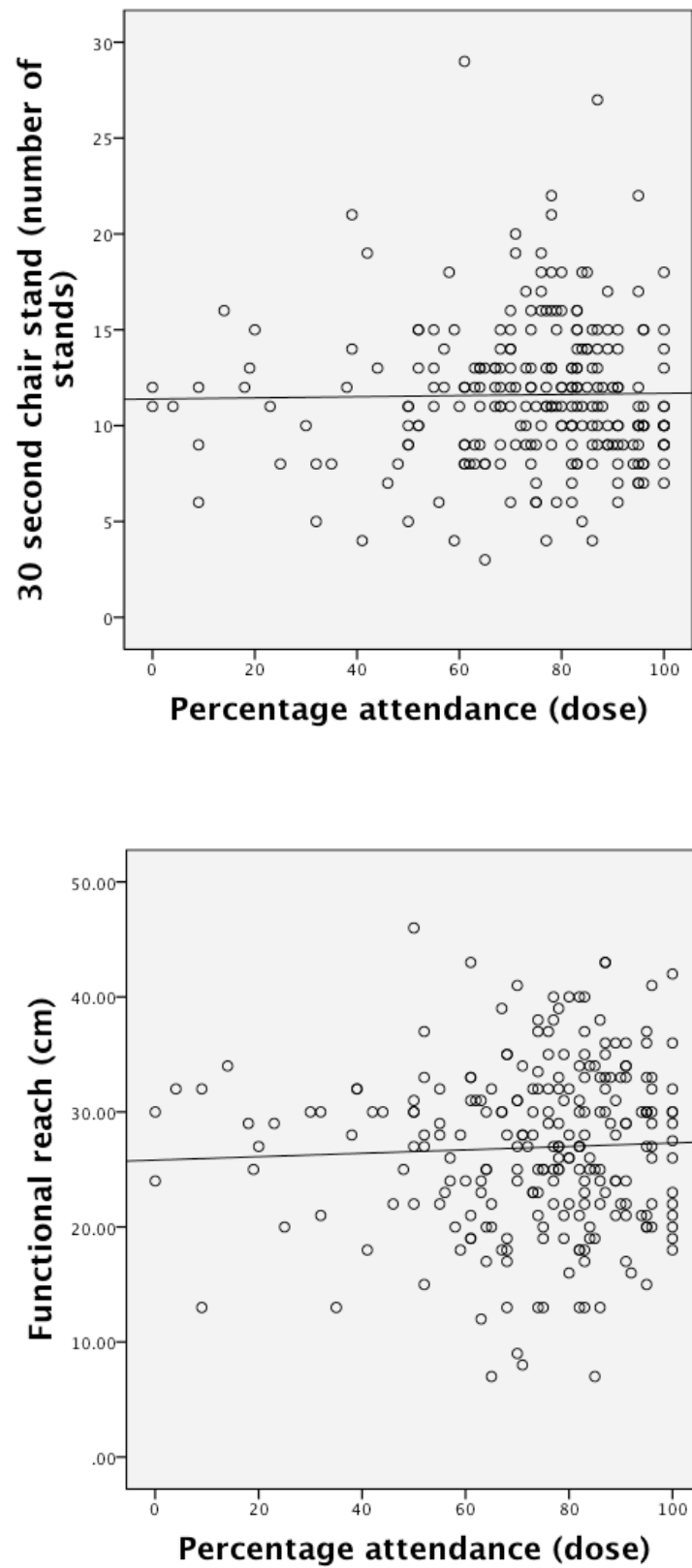
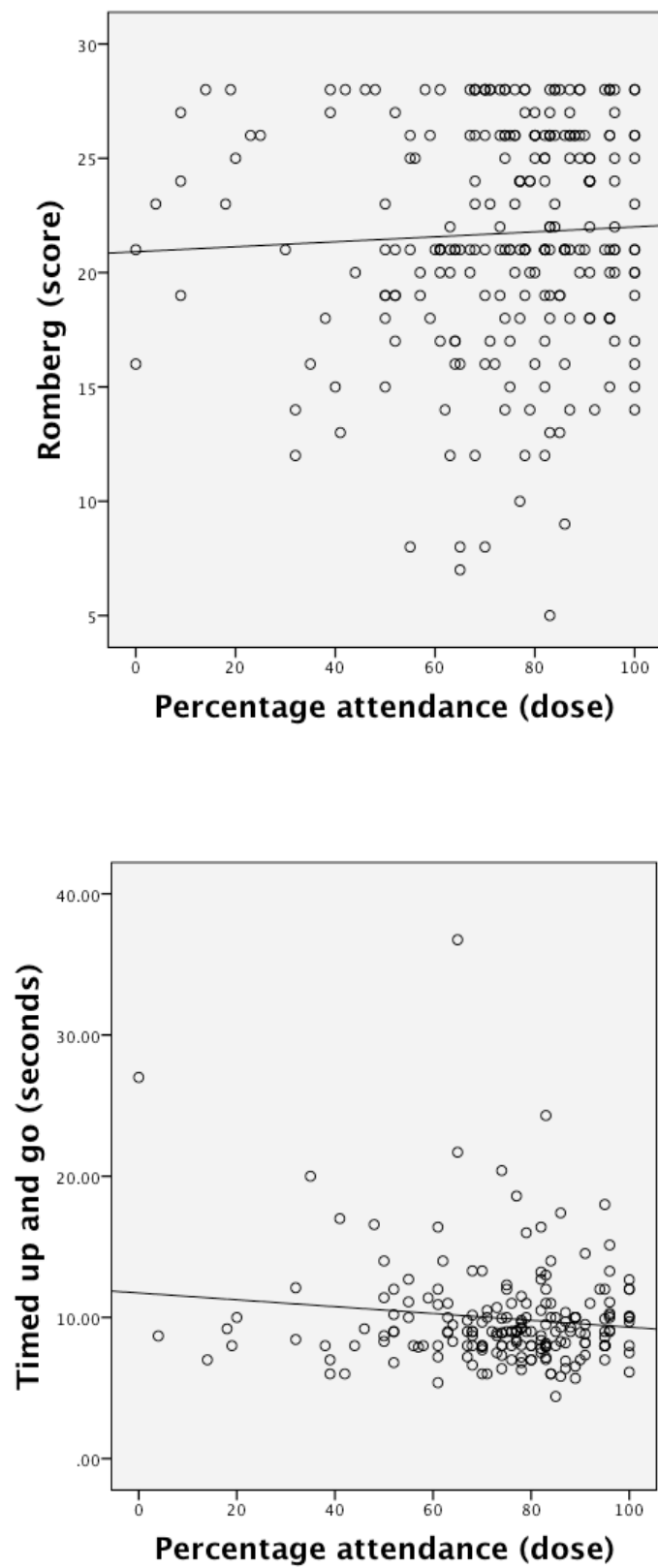


Figure 5.4, continued



Pearson bivariate correlation with 2-tailed significance test was used to further explore potential the relationship between dose and falls outcomes. The correlation coefficient measures the direction and strength of any linear relationship between the variables. Results are shown in Table 5.9; the correlation analyses exploring dose revealed no significant associations between paired variables.

**Table 5.9,** Correlation analyses outcomes for exercise dose (patient attendance)

Predictor variable	Dependent variable	Correlation coefficient	p value
Dose	Falls rate during the intervention	-0.052	0.385
	Falls rate during the 12 months post-intervention	-0.111	0.095
	Falls rate during the intervention and 12 months post-intervention	-0.080	0.180
	Timed up and go	-0.094	0.186
	Romberg balance	0.047	0.458
	Timed chair rise	0.009	0.889
	Functional reach	0.047	0.458

#### 5.6.2.1.2 Therapist Characteristics; PSI age & PSI attendance

Two of the three scatter plots in Figure 5.5 (falls during the intervention and falls during the combined intervention and 12 months post-intervention) show a weak positive association; falls incidence rate increases with older therapists. However, as discussed in relation to dose, there are many patients without any falls irrespective of therapist age, therefore, falls rate was dichotomised and histograms were produced (see Figure 5.6). There does not appear to be any clear association between falls and the age of the therapist, although it is interesting that older therapists have a higher proportion of fallers during the intervention but a lower proportion (than the other two age groups) in the 12 months post-intervention. There does not appear to be any effect of therapist age on the functional assessment outcomes (Figure 5.7).



**Figure 5.5, Correlation analyses for PSI age and falls rate**

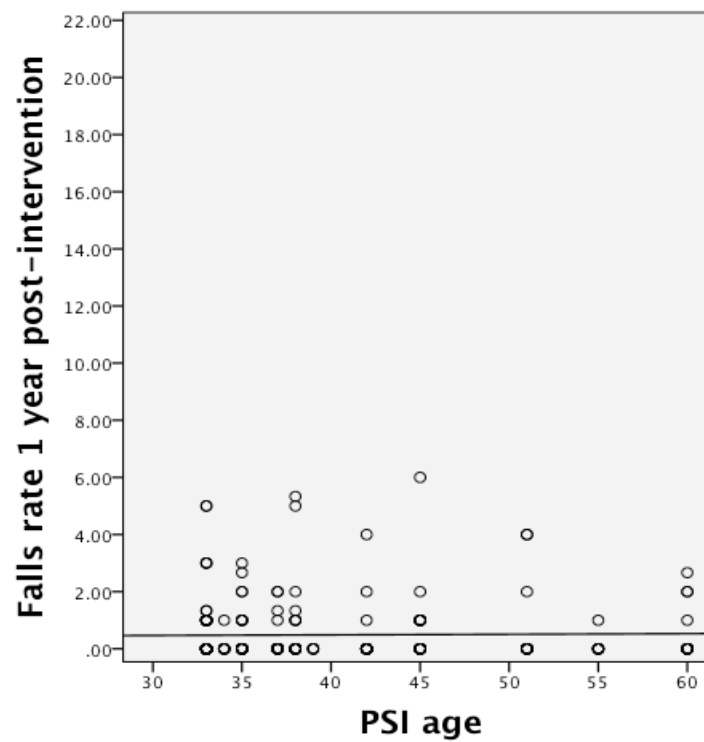
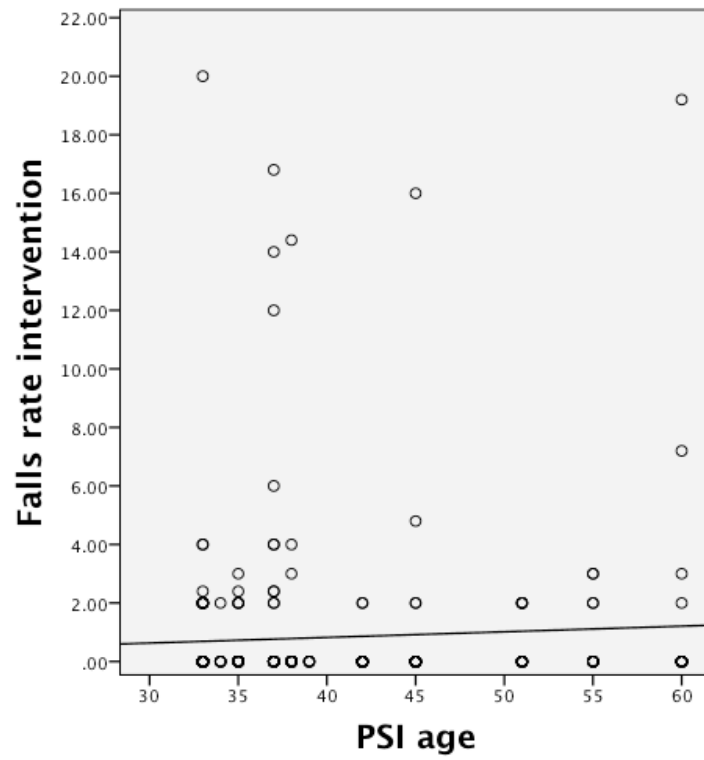
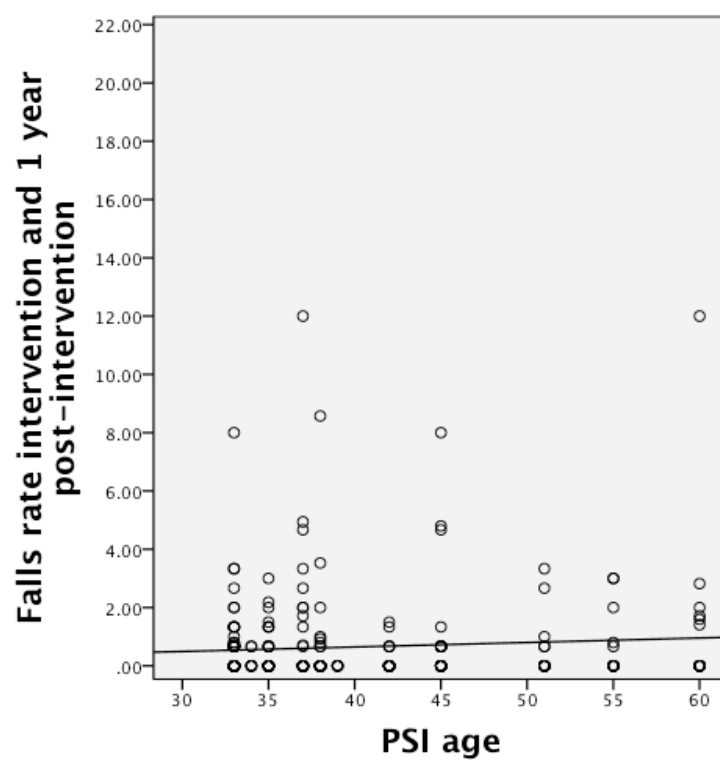
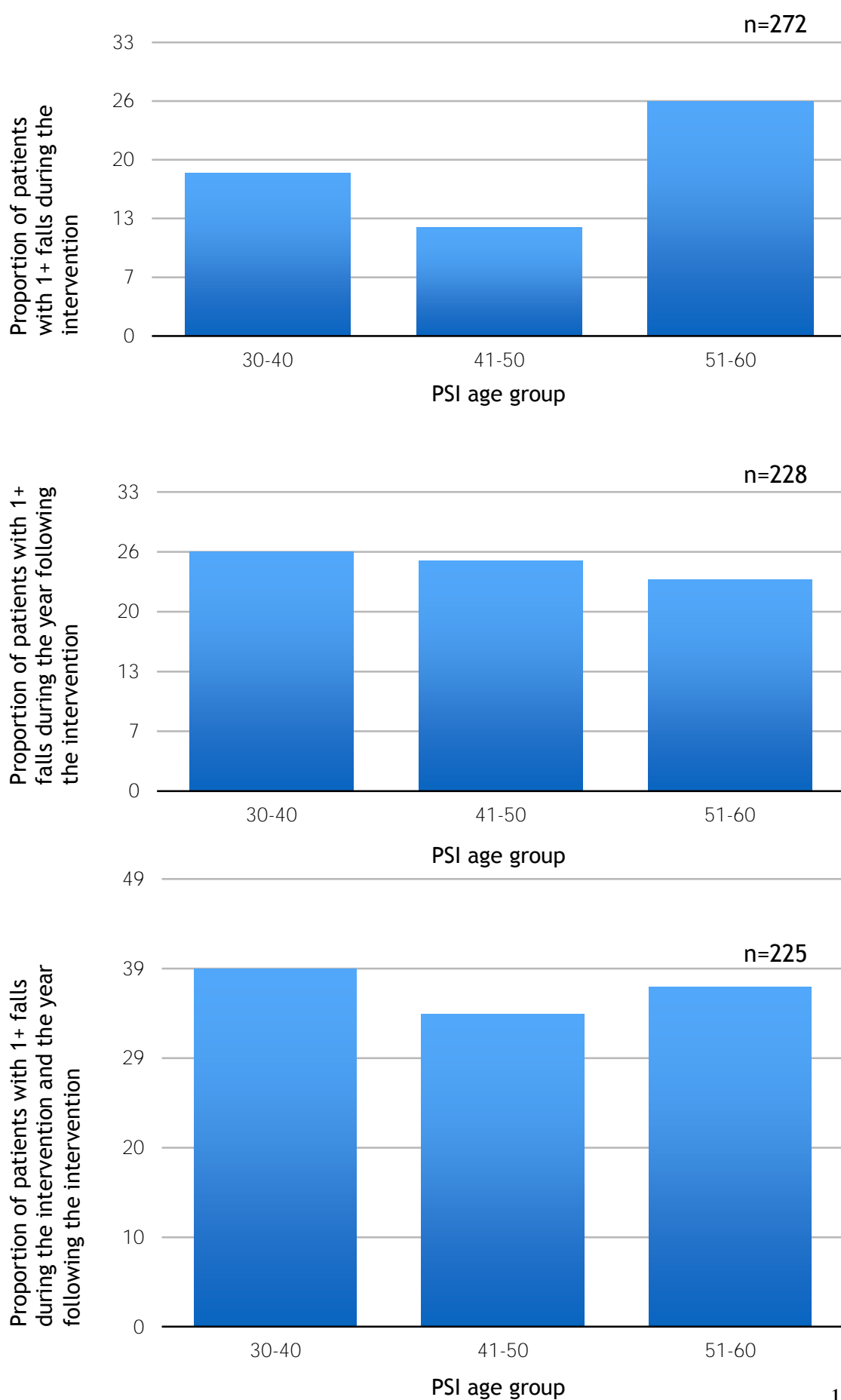


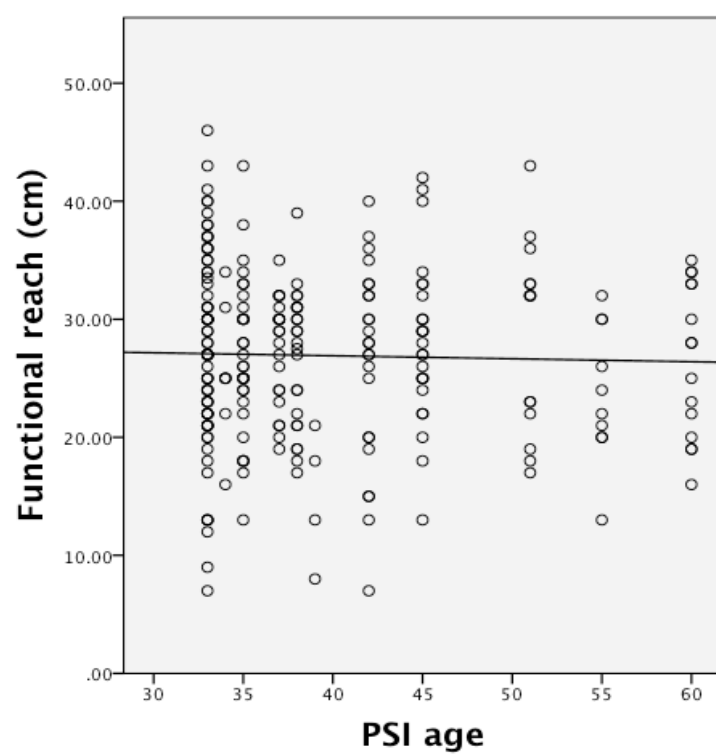
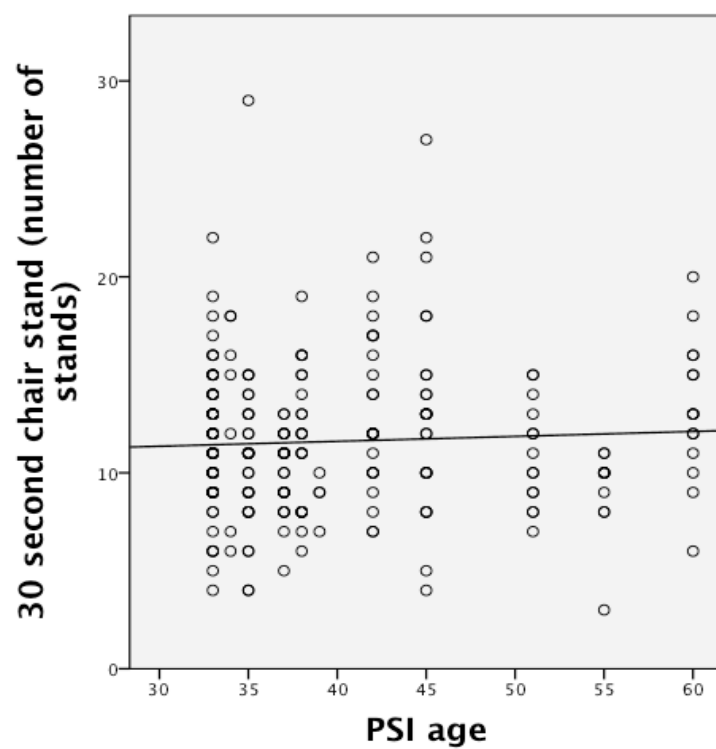
Figure 5.5, continued



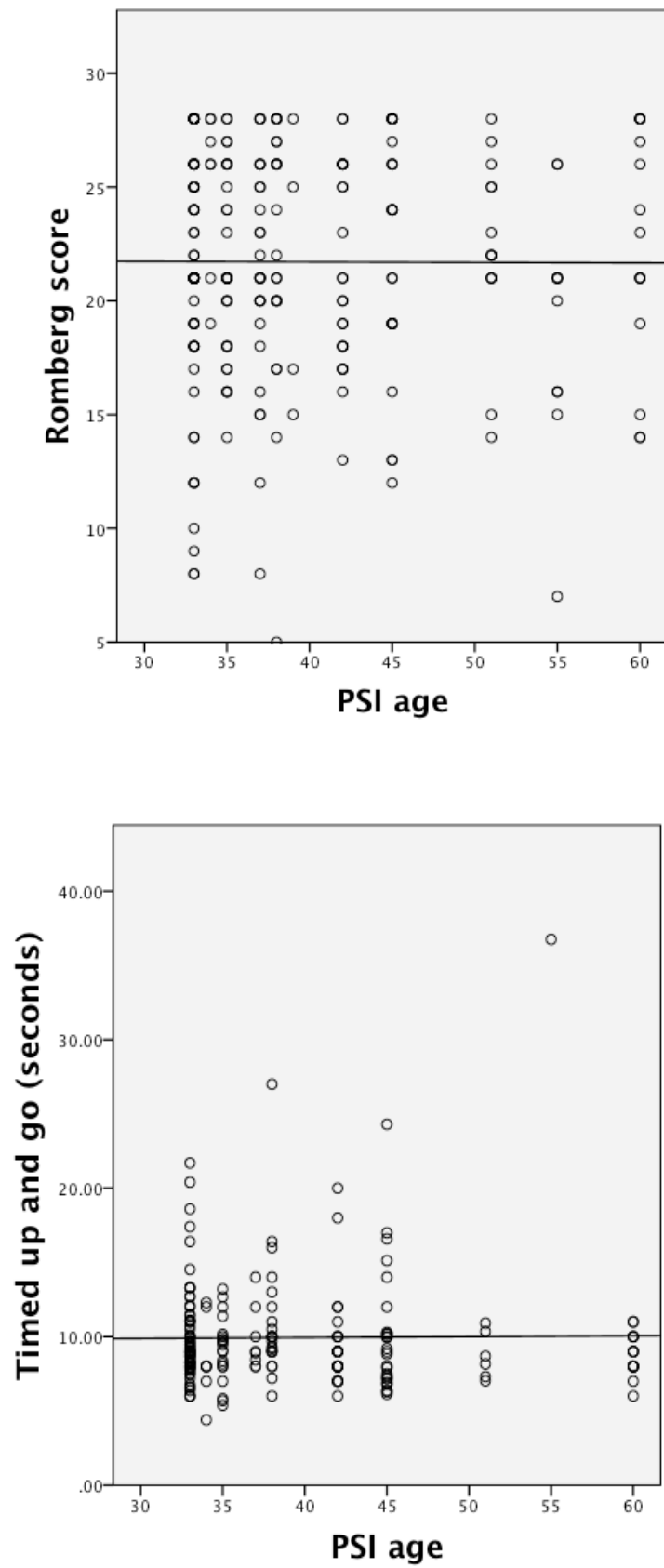
**Figure 5.6, Proportion of fallers (1+ falls) by PSI age and time-point**



**Figure 5.7,** Correlation analyses for PSI age and post-intervention functional outcomes



**Figure 5.7, continued**



For PSI attendance (the percentage of the exercise sessions they taught themselves), two of the three scatter plots shown in Figure 5.8 (falls during the intervention and falls during the combined intervention and 12 months post-intervention) show a weak positive association; falls incidence rate increases with therapists with higher attendance. The simpler histograms are shown in Figure 5.9. The proportion of fallers appears higher across all time-points when the intervention was consistently delivered by the known, allocated PSI.

Correlation analyses for PSI attendance and post-intervention functional outcomes are shown in Figure 5.10. As the percentage attendance of the allocated PSI increases, Romberg balance score decreases and timed-up-and-go increases, indicating that leg power, gait and balance is poorer. However, those patients whose instructor had higher attendance, also had improved functional reach; another measure of balance, so the results appear contradictory. There is no apparent association between PSI attendance and the chair rise outcome (Figure 5.10).

**Figure 5.8,** Correlation analyses for PSI attendance and falls rate

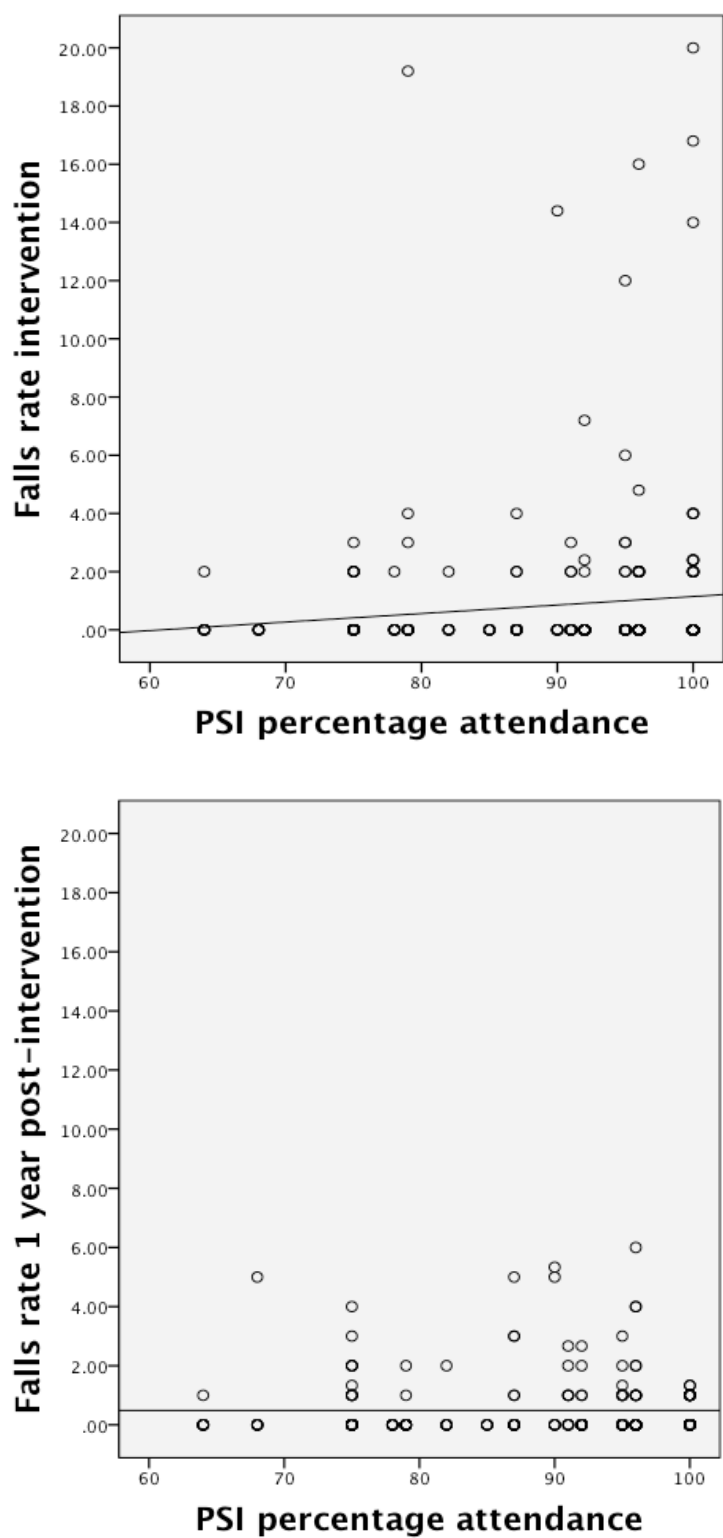
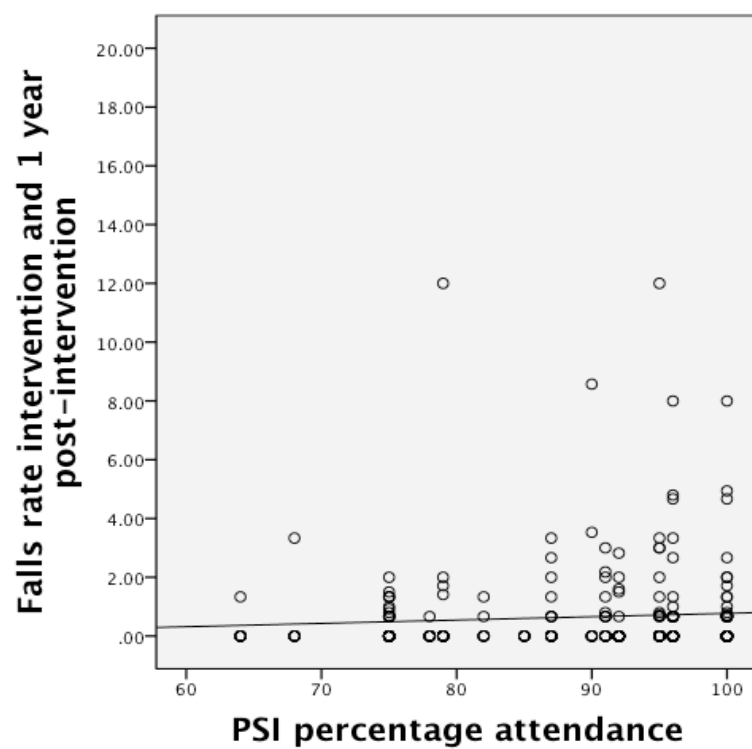
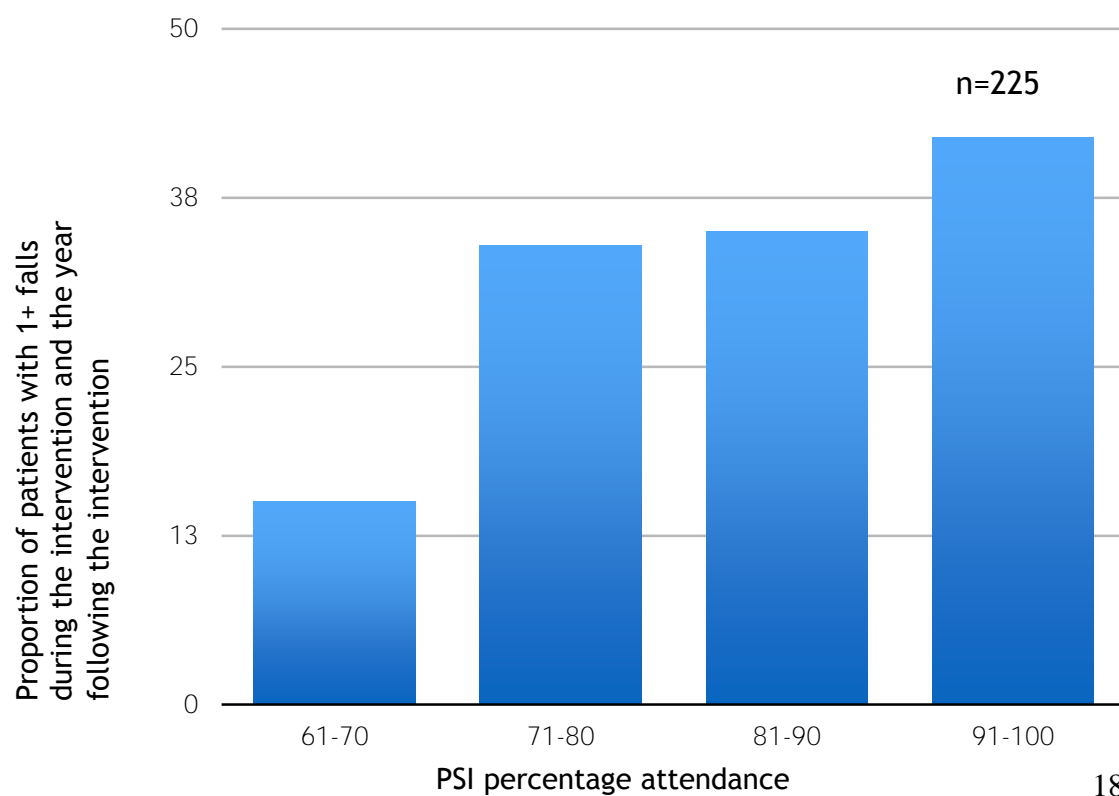
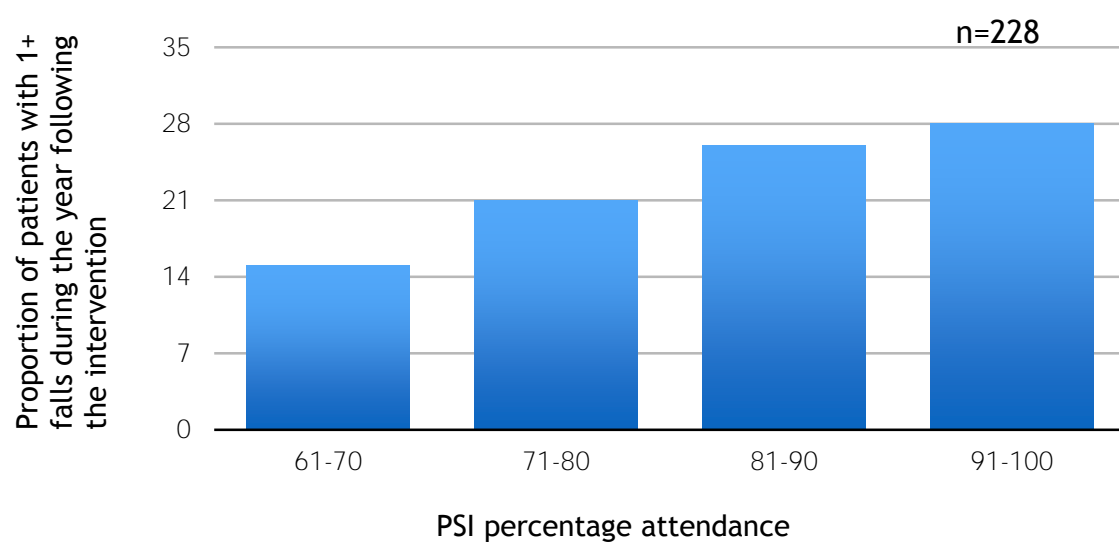
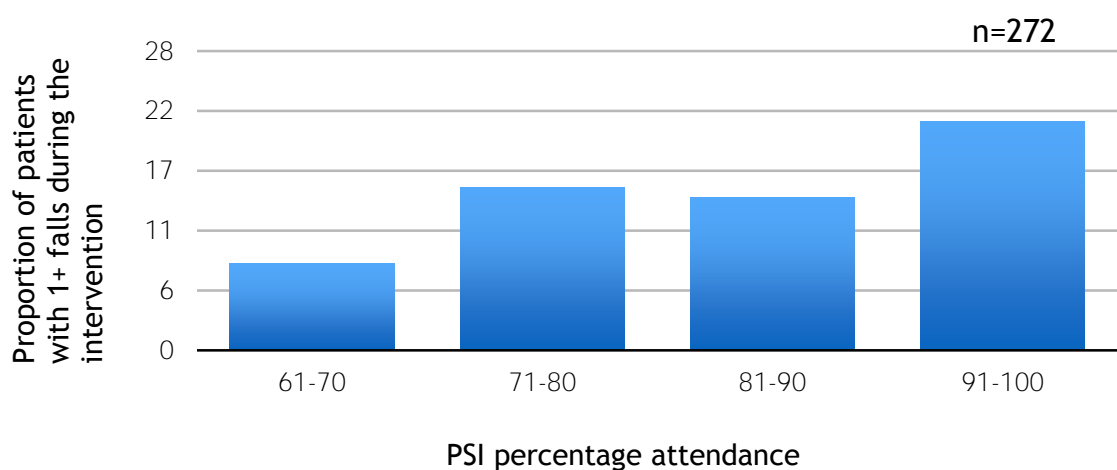


Figure 5.8, continued

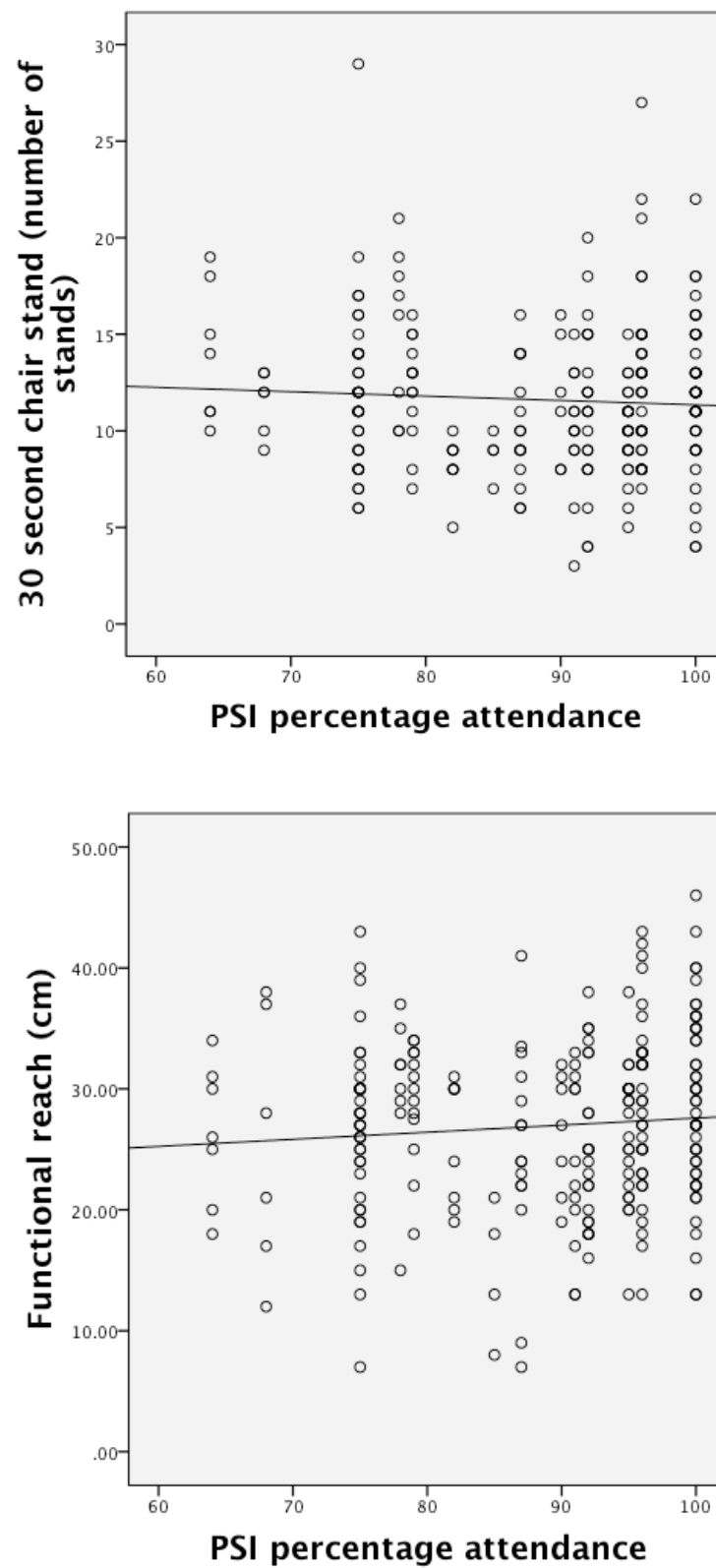




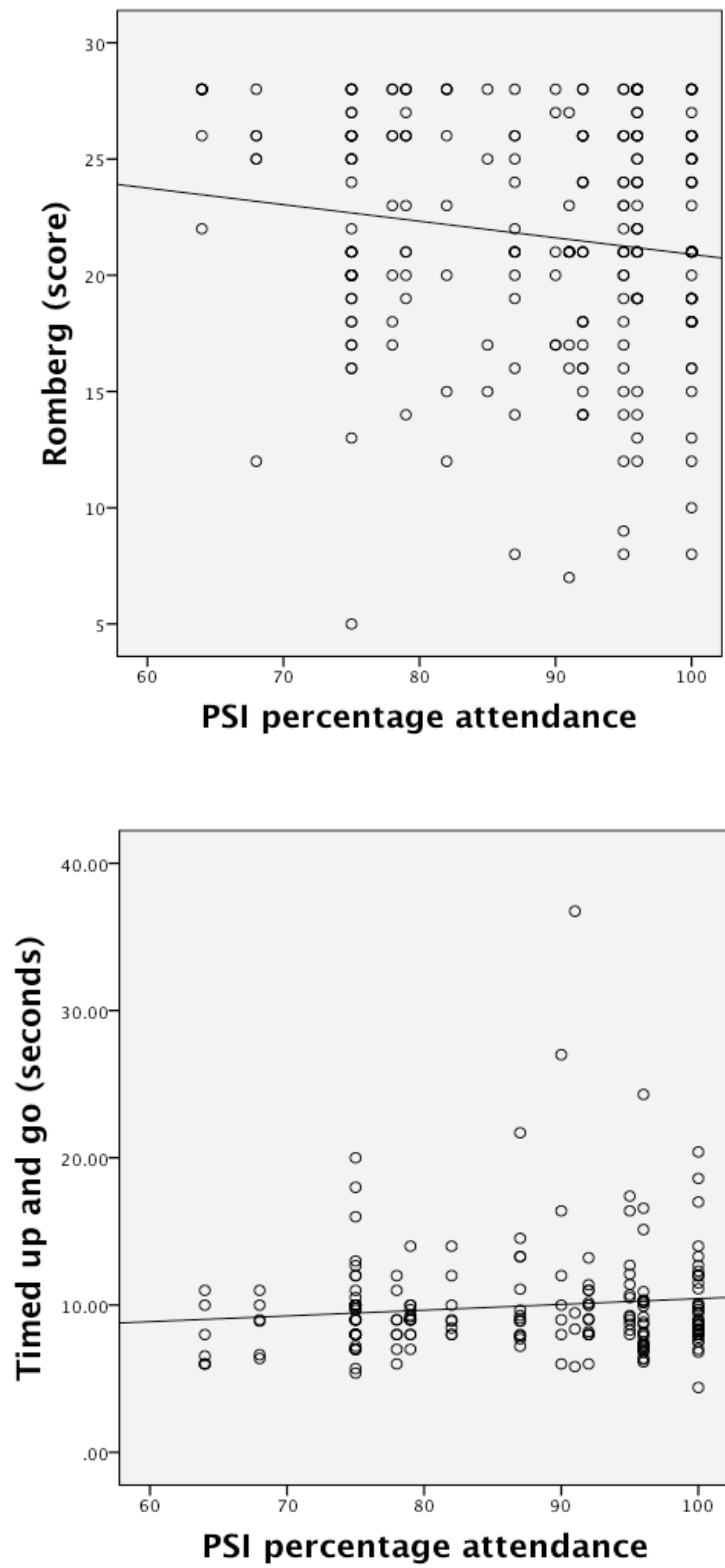
**Figure 5.9, Proportion of fallers (1+ falls) by PSI attendance and time-point**



**Figure 5.10,** Correlation analyses for PSI attendance and post-intervention functional outcomes



**Figure 5.10, Correlation**



Pearson bivariate correlation with 2-tailed significance test was used to further explore potential relationships between these two therapist characteristics (PSI attendance and age) and falls outcomes (Table 5.10). There was a significant ( $p<0.05$ ) association between only PSI attendance and post-intervention Romberg balance score. The Pearson correlation coefficient for these variables is -0.148, which is significant for a two-tailed test, based on a sample of 200 cases. The correlation was negative; Romberg balance score decreases as the percentage of classes the allocated (main) PSI taught themselves increases. The scatter plot for this association confirms the negative correlation (Figure 5.10), however, the associated R-squared value of 0.022 suggests that the model does not fit the data well. The correlation analyses exploring PSI age and PSI attendance revealed no other significant associations between paired variables (Table 5.10).

**Table 5.10**, Correlation analyses outcomes for PSI age & PSI attendance

Predictor variable	Dependent variable	Correlation coefficient	p value
PSI Age	Falls rate during the intervention	0.056	0.354
	Falls rate during the 12 months post-intervention	0.015	0.820
	Falls rate during the intervention and 12 months post-intervention	0.079	0.193
	Timed up and go	0.013	0.860
	Romberg balance	-0.003	0.956
	Timed chair rise	0.053	0.408
	Functional reach	-0.028	0.667
PSI Attendance	Falls rate during the intervention	0.108	0.075
	Falls rate during the 12 months post-intervention	0.000	0.996
	Falls rate during the intervention and 12 months post-intervention	0.073	0.227
	Timed up and go	0.112	0.116
	Romberg balance	-0.148	0.020
	Timed chair rise	-0.062	0.335
	Functional reach	0.083	0.198

### 5.6.2.2 T-test analyses; therapist characteristics

The t-tests reported in this section refer to the descriptive statistics given in Table 5.7. Independent samples t-test analysis was used to explore the differences in mean patient outcome scores by the following PSI characteristics; gender, professional background, clinical experience and the timing of PSI training. T-test analyses outcomes are shown in Table 5.11. When the mean number of chair stands was compared, patients whose PSI was trained during the trial had a statistically significantly higher mean number of chair stands ( $12.7 \pm 4.2$  stands) at the end of the exercise intervention compared to those whose PSI was trained before the trial ( $11.0 \pm 3.4$  stands),  $t(244) = -3.545$ ,  $p = <0.001$ . However, this analysis was exploratory, and, due to multiple testing, there was a high risk of chance finding. The t-tests revealed no other significant differences between compared means.

**Table 5.11**, T-test scores and p values

Predictor variable	Dependent variable	t score (df)	p value
PSI Gender	Falls rate during the intervention	$t(270) = -0.385$	0.700
	Falls rate during the 12 months post-intervention	$t(273) = -0.411$	0.682
	Falls rate during the intervention and 12 months post-intervention	$t(226) = -1.241$	0.216
	Timed up and go	$t(196) = 1.629$	0.105
	Functional reach	$t(241) = -1.440$	0.151
	Romberg balance	$t(246) = -0.285$	0.776
	Timed chair rise	$t(244) = -1.096$	0.274
PSI Professional background (physiotherapist or exercise professional)	Falls rate during the intervention	$t(270) = 0.100$	0.920
	Falls rate during the 12 months post-intervention	$t(273) = 0.078$	0.938
	Falls rate during the intervention and 12 months post-intervention	$t(226) = -0.759$	0.449
	Timed up and go	$t(196) = -0.784$	0.434
	Functional reach	$t(241) = 0.397$	0.692
	Romberg balance	$t(246) = -1.455$	0.147

Predictor variable	Dependent variable	t score (df)	p value
	Timed chair rise	t(244)= -0.755	0.451
PSI Clinical experience	Falls rate during the intervention	t(270)= 1.043	0.298
	Falls rate during the 12 months post-intervention	t(273)= 0.553	0.580
	Falls rate during the intervention and 12 months post-intervention	t(226)= -0.875	0.382
	Timed up and go	t(196)= -0.742	0.459
	Functional reach	t(241)= -0.461	0.645
	Romberg balance	t(246)= 1.038	0.300
	Timed chair rise	t(244)= -1.818	0.070
PSI Training (trained during or before the trial)	Falls rate during the intervention	t(270)= -0.257	0.797
	Falls rate during the 12 months post-intervention	t(273)= -0.674	0.501
	Falls rate during the intervention and 12 months post-intervention	t(226)= -0.826	0.410
	Timed up and go	t(196)= -0.196	0.845
	Functional reach	t(241)= -0.246	0.806
	Romberg balance	t(246)= -0.914	0.362
	Timed chair rise	t(244)= -3.545	<0.001

### 5.6.2.3 Analysis of variance (ANOVA); therapist characteristics

The ANOVA analyses reported in this section refer to the descriptive statistics given in Tables 5.7 and 5.8. I conducted one-way ANOVAs to explore the differences in mean patient outcome scores by the PSI characteristics shown in Table 5.12. One-way ANOVA revealed a statistically significant difference between groups of patients according to their enjoyment of the exercise classes for falls rates at all time points and for functional reach, as well as between groups of patients according to their rating of exercise class intensity for all functional assessments. The F-values, which indicate the

percentage variance explained, and between group significance values are shown in Table 5.12.

**Table 5.12**, ANOVA scores and p values

Predictor variable	Dependent variable	F value (df)	p value
PSI 'Quality' (low, moderate or high)	Falls rate during the intervention	F(2,269) = 1.541	0.216
	Falls rate during the 12 months post-intervention	F(2,225) = 0.819	0.442
	Falls rate during the intervention and 12 months post-intervention	F(2,272) = 1.206	0.301
	Timed up and go	F(2,195) = 1.845	0.161
	Functional reach	F(2,240) = 0.827	0.439
	Romberg balance	F(2,245) = 0.417	0.660
	Timed chair rise	F(2,243) = 2.340	0.099
Patient enjoyment of classes (Far too easy, A little too easy, About right, A little too hard, Far too hard)	Falls rate during the intervention	F(3,189) = 4.758	0.003
	Falls rate during the 12 months post-intervention	F(3,180) = 5.197	0.002
	Falls rate during the intervention and 12 months post-intervention	F(3,190) = 6.268	<0.001
	Timed up and go	F(3,148) = 0.150	0.930
	Functional reach	F(3,178) = 3.079	0.029
	Romberg balance	F(3,181) = 0.786	0.503
	Timed chair rise	F(3,180) = 0.416	0.742
Patient-rated intensity of classes (Very enjoyable, Somewhat enjoyable, Neither enjoyable nor not enjoyable, Not really enjoyable, Not enjoyable at all)	Falls rate during the intervention	F(3,191) = 0.156	0.926
	Falls rate during the 12 months post-intervention	F(3,181) = 0.372	0.774
	Falls rate during the intervention and 12 months post-intervention	F(3,191) = 0.115	0.951
	Timed up and go	F(3,149) = 6.383	<0.001
	Functional reach	F(3,179) = 6.392	<0.001
	Romberg balance	F(3,182) = 5.301	0.002
	Timed chair rise	F(3,181) = 11.236	<0.001

#### **5.6.2.3.1 Patient enjoyment of classes**

Tukey post-hoc tests (conducted as part of the ANOVA analyses) showed that falls rate during the intervention was statistically significantly lower in the group of patients who thought the classes were 'very enjoyable' ( $0.6 \pm (\text{standard deviation}) 2.0$ ,  $p=0.002$ ) and 'somewhat enjoyable' ( $0.5 \pm 1.0$ ,  $p=0.002$ ) compared to the group who stated the exercise was 'neither enjoyable nor not enjoyable' ( $4.0 \pm 6.2$ ). The difference in falls rate during the intervention between the 'very enjoyable' and 'somewhat enjoyable' groups was not statistically significant ( $p=0.980$ ). The falls rate in the intervention and 12 months following the intervention was also statistically significantly lower in the group of patients who thought the classes were 'very enjoyable' ( $0.5 \pm 1.0$ ,  $p=0.010$ ) and 'somewhat enjoyable' ( $0.6 \pm 0.9$ ,  $p=0.033$ ) compared to the group who stated the exercise was 'neither enjoyable nor not enjoyable' ( $1.9 \pm 2.1$ ). Mean functional reach in centimetres was significantly poorer in the group of patients who thought the classes were 'neither enjoyable nor not enjoyable' ( $17.5 \pm 2.1$  cm) compared with those who thought the classes were 'very enjoyable' ( $27.0 \pm 7.3$  cm) and 'somewhat enjoyable' ( $28.2 \pm 7.1$  cm).

Enjoyment of classes was strongly associated with reduced falls at all time-points (Table 5.12). Intervention dose, however, was not associated with reduced falls (Table 5.10). This suggests that enjoyment of classes was not associated with better adherence to the intervention, therefore an additional ANOVA analysis was conducted to explore this, looking at the direct effect of enjoyment of classes on class attendance (dose). This revealed a statistically significant difference in mean exercise dose between groups of patients according to their enjoyment of the exercise classes ( $F(3,195) = 5.484$ ,  $p=0.001$ ). Tukey post-hoc tests showed that percentage attendance was statistically significantly higher in the group of patients who thought the classes were 'very enjoyable' compared to the group who stated the exercise was 'neither enjoyable nor not enjoyable' ( $5.1 \pm 49.9$ ,  $p=0.009$ ).



### 5.6.2.3.2. Patient-rated intensity of classes

Mean and standard deviations for functional assessments scores between patient-rated intensity groups are shown in Table 5.13. Tukey post-hoc tests for patients grouped according to their rating of exercise intensity showed that:

- Timed up and go was significantly faster in the groups who thought the classes were 'far too easy' ( $p=0.006$ ) and 'a little too easy' ( $p=0.004$ ) compared with those who said the sessions were 'a little too hard'.
- Functional reach was significantly shorter, indicating poorer balance, in all groups compared with the 'far too easy' group ('a little too easy'  $p=0.004$ , 'about right'  $p=0.001$ , 'a little too hard'  $p=0.001$ ).
- Romberg balance score was significantly better in the 'far too easy' group ( $p=0.007$ ) and the 'a little too easy' group ( $p=0.013$ ) compared with those who said the sessions were 'a little too hard'.
- Significantly more chair stands were performed in 30 seconds in all groups compared with those who thought the exercise was too hard ('far too easy'  $p<0.001$ , 'a little too easy'  $p=0.001$ , 'about right'  $p=0.028$ ).

**Table 5.13**, Mean and standard deviation for functional assessments scores between patient-rated intensity groups

Functional assessment	Group				
	Far too easy	A little too easy	About right	A little too hard	Far too hard
TUG (seconds)	7.9 ± 1.9	8.3 ± 1.6	10.1 ± 3.3	13.4 ± 3.4	Insufficient cases to include in the analysis
Functional reach (cm)	35.9 ± 4.0	27.0 ± 7.2	26.7 ± 7.3	21.8 ± 5.6	
Romberg score	25.2 ± 2.6	23.6 ± 4.0	21.5 ± 4.9	17.0 ± 8.5	
Chair rise (stands)	15.9 ± 4.9	13.3 ± 3.7	11.3 ± 3.3	7.3 ± 2.2	

### 5.6.2.4 Summary; bivariate analyses

Chair rise group means were non-significantly associated with two predictor variables (pre-trial experience of delivering FaME and PSI 'quality' rating) and are significantly associated with the timing of PSI training and patients' enjoyment of the classes (Table 5.14). Some results are unexpected, for example, greater leg power was achieved by

inexperienced PSIs and PSIs who achieved a low quality assurance rating. This needs further investigation and suggests that adjusting for baseline functional assessment score is needed when carrying out multilevel modelling (see section 5.6.4.2).

**Table 5.14,** Chair rise outcomes

Group	n	Mean number of chair stands (SD)	Significance
Previous FaME teaching experience	70	10.91 (3.02)	p=0.07
No previous FaME teaching experience	176	11.88 (4.01)	
PSI trained before trial	160	10.99 (3.41)	p<0.001
PSI trained during the trial	86	12.74 (4.16)	
Low PSI quality assurance (QA) rating	25	13.04 (5.24)	p=0.09 for low quality compared with moderate
Moderate PSI QA rating	82	11.18 (3.81)	
High PSI QA rating	139	11.60 (3.40)	

Enjoyment of classes was strongly associated with reduced falls at all time-points. Intervention dose, however, was not associated with reduced falls. An additional analysis revealed a difference in mean exercise dose between groups of patients according to their enjoyment of the exercise classes. This needs further investigation and may be elucidated by the multi-level modelling analyses.

Patient-rated intensity of classes was strongly associated with all falls risk factors (indicated by functional assessment scores) but not with falls at any time-point. Multi-level modelling analyses should therefore adjust for baseline functional assessment score.

Consistent attendance of the allocated (known) therapist was negatively associated with Romberg balance score. Adjusting for baseline functional assessment score in the multi-level modelling analyses may help resolve this unexpected finding.

### **5.6.3 Multiple regression equations**

Multiple regression equations were planned to adjust for any confounding between therapist characteristics that had a positive effect on falls rate or falls risk factors. However, as there was no more than one therapist characteristic that appeared to be related to a particular patient outcome, multiple regression was deemed unnecessary.

### **5.6.4 Multilevel modelling**

Multilevel modelling (MLM) was used to allow for clustering of patients within therapists and estimated variances at each of two levels; the lower, patient level and the higher, therapist level. This allowed me to investigate the extent of grouping in patient outcomes by exercise therapist (PSI), as well as to investigate possible predictors of grouping in patient outcomes by PSI characteristics.

#### **5.6.4.1 Unconditional models**

Unconditional models are models without predictors. These were used to identify any therapist effect for the selected range of dependent variables (falls rates and falls risk factors). The unconditional models are shown in Table 5.15. SPSS was unable to process the data for three models (falls rate in the 12 months post-intervention, falls rate during the intervention and the 12 months post-intervention and functional reach) due to the intercept covariance parameter being redundant. This indicates that there was no evidence of between therapist variation, and therefore further exploration using multi-level modelling is also redundant since there is no therapist variation for which to find an explanatory, predictor variable. For all models shown in Table 5.15, with the exception of TUG, the intraclass correlation coefficient (ICC) is low, indicating that patients grouped by PSI lack similarity with regard to the dependent variable (falls rate during the intervention, Romberg and chair rise). Tests for the statistical significance of the variance for each of these dependent variables give p-values greater than 0.05, which, along with the associated low ICC, would not traditionally support the use of multilevel modelling. The TUG ICC, however, was very high (0.81) indicating that patients grouped by PSI correlate strongly; 81% of the variance is at the group (PSI) level. TUG is further discussed later in section 5.6.4.2.1. The unconditional models lack robust evidence of a therapist effect and the difference

in means can mainly be attributed to random variation arising from patients. However, any ICC of greater than zero indicates that some variance is not explained by differences between individuals. In addition, given that the trial attempted to standardise the delivery of the exercise intervention (via PSI trial training and quality assurance visits), large variation between therapists would be unlikely. Moreover, therapist effects are known to be small in psychotherapy. These factors provided a rationale for the addition of predictor variables to the models in order to explore causality of variance.

**Table 5.15**, Multilevel modelling; unconditional models

Dependent Variable	Deviance (-2LL)	Variance within PSIs, between patients	Variance between PSIs	ICC	p-value for PSI variance
Falls rate during the intervention	1318	7.35	0.11	0.01	0.54
Romberg	1496	23.96	0.54	0.02	0.51
Chair rise	1347	13.46	0.89	0.06	0.23
TUG	1071	10.52	43.47	0.81	0.06

#### 5.6.4.2 Adding predictors to the models

There was no evidence of any between therapist variation for three dependent variables (falls rate in the 12 months post-intervention, falls rate in the combined intervention and 12 months post-intervention and functional reach), so further exploration was only necessary for the four remaining outcomes (falls rate during the intervention, Romberg balance, TUG and Chair rise).

The ten proposed predictor variables (used in the bivariate analyses) were potentially too many, with some being too similar, therefore the number used in the MLM analysis was reduced. The bivariate analyses suggested that professional background was not associated with change in any of the dependent variables (DVs), plus the number of physiotherapists was small (2 of 12 PSIs). All other predictor variables (PVs) were found to have at least a weak association with one or more DVs, supporting their

retention in the more complex analyses. However, I decided that the timing of the achievement of the PSI qualification was directly linked to achievement of pre-trial clinical experience of FaME intervention delivery, therefore, the 'training' PV was also removed from the analysis. This left eight PVs.

In addition, the baseline score of the appropriate outcome variable was added to the models to control for intrinsic differences between patients prior to the intervention. Introducing the baseline score was only possible for the three functional assessment outcomes (Romberg, TUG and chair rise) for which a baseline score was taken.

As reported in section 5.4 data from patient evaluation questionnaires were only available for 196 patients in the TDE group. When these data were introduced to the multilevel models, there was a significant reduction in deviance that resulted from the smaller sample size. As change in deviance was being used to evaluate the effect of fitting the predictor variable, the smaller sample size impacted on the analysis. For this reason, additional unconditional and adjusted models were produced using data for only the 196 patients who had returned their evaluations. The deviance from these models was used as the comparator to calculate any change in deviance when the two variables from the evaluation were fitted. Tables presented for all multilevel models therefore show the results for the full TDE sample and for the smaller (196 patients) sample.

As stated above, change in deviance was used to evaluate the effect of fitting a predictor variable. The difference in deviance (or -2xLog-Likelihood ratio (-2LL)) between the two models, one with the variable fitted, one without, was calculated, and the effect of the variable was tested for significance using a separate chi-square test. This is considered the best way to test the effect of the variable according to statistical orthodoxy (Goldstein, 2011).

#### **5.6.4.2.1 Timed Up & Go (TUG)**

The TUG unconditional model had an intra-class correlation (ICC) of 0.81 indicating that patients grouped by PSI correlate strongly; 81% of the variance was at the group

(PSI) level. The associated test for statistical significance of the variance (43.47) was  $p=0.06$ , 95% CI 15, 123. Adjusting for baseline TUG score improves the model significantly, but the corresponding ICC shows that once the baseline score has been controlled for, only 0.3% of the variance is explained by patients grouped by PSI (Table 5.16). This was surprising considering that 81% of the variance appeared to be at the PSI level in the unconditional model. In order to explore this, mean TUG and standard deviations for each PSI were investigated by producing simple means tables in SPSS (Table 5.17). Some missing TUG scores were caused by lack of availability of sufficient space to perform the assessment in the general practice. As patients were allocated to study group by GP practice, this resulted in missing TUG scores being clustered by PSI. Table 5.17 shows that all baseline TUG tests were missing for PSI06 and all follow up TUG tests were missing for PSI04. In addition, PSI05 has only 1 follow up TUG score of 36.75 seconds and this has had the effect of skewing the mean change in TUG for this PSI. For these reasons, data for PSIs 04, 05 & 06 were removed from the MLM analysis as a post-hoc, investigative measure and the results are shown in Table 5.18. This reduced the variance at group (PSI) level to 2% in the unconditional model. Figure 5.11 shows box plots of TUG scores before and after the intervention, as well as the change in TUG, with all data and PSIs included. The outlier associated with PSI05 can clearly be seen in all plots.

**Table 5.16, Multilevel modelling; TUG**

Model	Deviance (-2LL), df	Change in deviance, df	Is change in deviance significant ?	Effect of variable p-value	Patient level residual variance	PSI level intercept variance	ICC
Unconditional	1071, 1	-	-	-	10.52	43.47	0.810
Unconditional (n=196)	785, 1	-	-	-	9.72	0.22	0.022
Baseline TUG	778, 2	293, 1	Y	<0.001	4.20	0.01	0.003
Baseline TUG (n=196)	598, 2	187, 1	Y	<0.001	4.18	0*	-
Baseline TUG, PSI age	773, 3	5, 1	Y	0.015	4.10	0*	-
Baseline TUG, PSI gender	778, 3	0, 1	N	-	4.20	0.14	0.032
Baseline TUG, patient-rated intensity	592, 5	6, 3	Y	0.049	4.03	0*	-
Baseline TUG, patient enjoyment	586, 5	12, 3	Y	0.003	3.98	0*	-
Baseline TUG, PSI quality	777, 4	1, 2	N	0.303	4.19	0*	-
Baseline TUG, PSI attendance	776, 3	2, 1	N	0.104	4.16	0*	-
Baseline TUG, dose	778, 3	0, 1	N	-	4.20	0.17	0.039
Baseline TUG, PSI clinical experience	778, 3	0, 1	N	-	4.21	0.01	0.002

0\* = therapist level variation indistinguishable from patient level variation

**Table 5.17, TUG mean and SD by PSI and time-point**

PSI ID number	N	Baseline TUG			Follow up TUG			Change in TUG		
		mean	N	SD	mean	N	SD	mean	N	SD
PSI01	8	8.12	8	1.87	8.56	7	2.78	-0.31	7	2.26
PSI02	15	8.61	15	3.10	8.11	13	1.96	-0.12	13	1.27
PSI03	36	9.05	34	1.92	9.79	8	2.15	-0.30	8	1.35
PSI04	9	11.89	8	4.27	-	-	-	-	-	-
PSI05	15	14.27	9	15.70	36.75	1	-	19.22	1	-
PSI06	17	-	-	-	8.74	6	1.59	-	-	-
PSI07	43	9.84	38	3.06	9.13	20	2.19	-0.46	19	2.10
PSI08	24	12.52	24	5.52	11.01	23	4.32	1.30	23	2.75
PSI09	28	10.45	28	2.81	9.58	24	3.30	0.50	24	1.86
PSI10	19	10.96	19	3.56	8.80	15	1.42	1.46	15	3.08
PSI11	29	12.00	25	8.35	10.17	25	4.30	1.51	21	5.28
PSI12	71	10.38	65	2.75	10.40	56	3.26	-0.09	51	2.24

**Table 5.18, Multilevel modelling; TUG**

Model	Deviance (-2LL), df	Change in deviance, df	Is change in deviance significant?	Patient level residual variance	PSI level intercept variance	ICC
Unconditional	996, 1	-	-	10.61	0.20	0.019
Baseline TUG	769, 2	227, 1	Y	4.11	0*	-

0\* = therapist level variation indistinguishable from patient level variation



**Figure 5.11, Box plots for TUG by PSI**

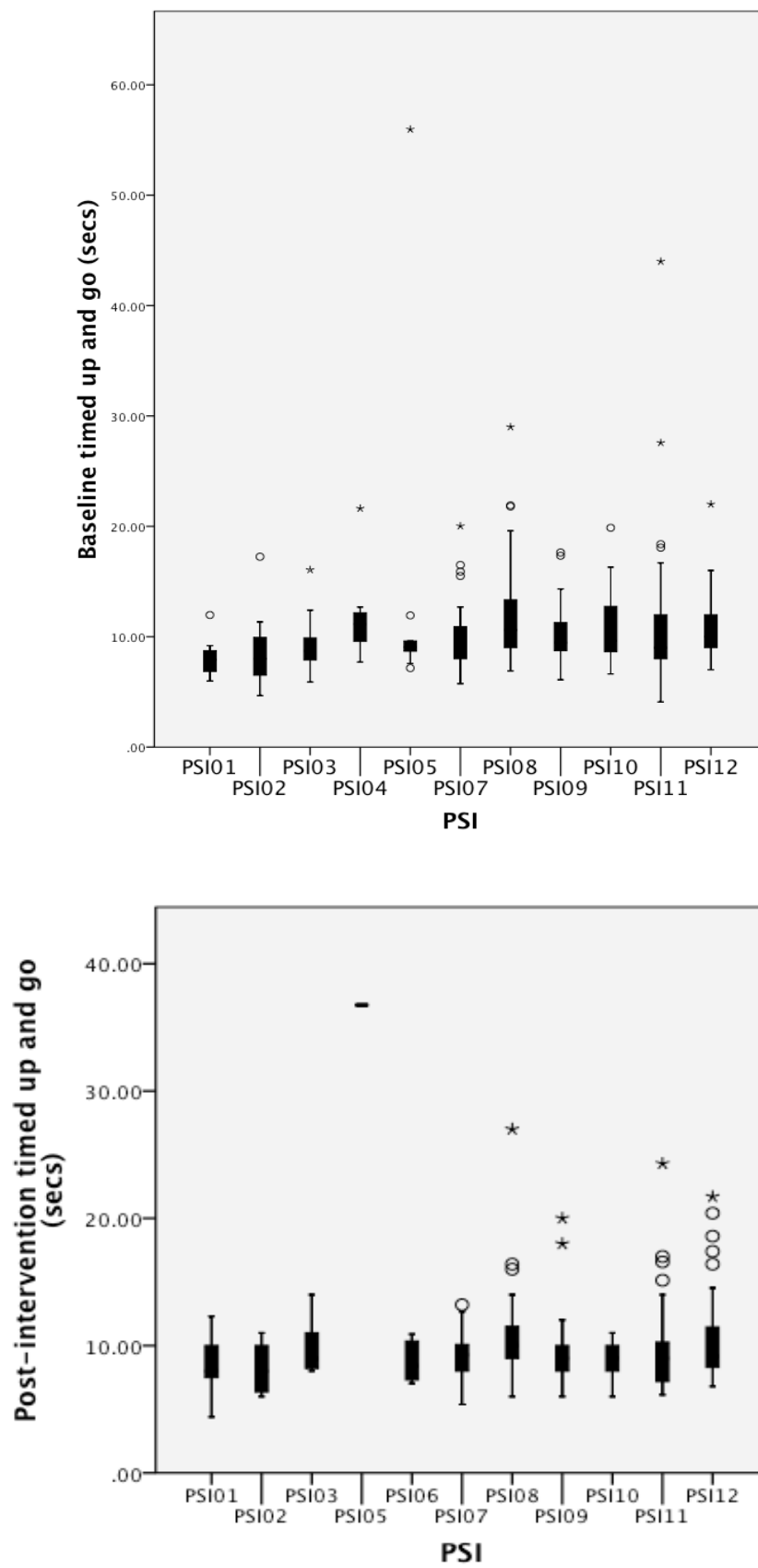
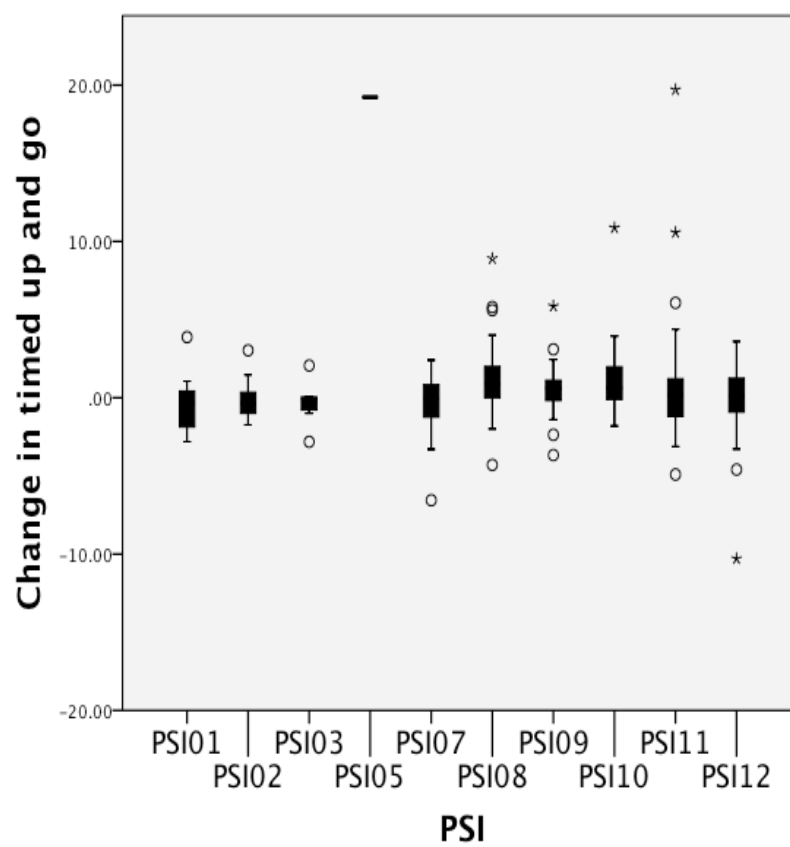


Figure 5.11, continued



When all data are included in the analysis, adjusting for baseline TUG score results in a significant change in deviance, which indicates that the variable added to the model (baseline TUG time) is an important predictor of post-intervention TUG time. Adding baseline TUG lowered the residual variance from 10.52 to 4.20, therefore, 60%  $((10.52 - 4.20)/10.52)$  of patient level variance has been explained by baseline TUG. PSI level variance lowered from 43.47 to 0.01, which indicated that 99% of PSI level variance has been explained by baseline TUG. Independent addition of five predictor variables to the adjusted model caused the intercept variance to become zero (Table 5.16). This indicated that there was no evidence of any between therapist variation once baseline score had been controlled for. The independent addition of the three remaining predictor variables (PSI gender, exercise dose and the experience of the PSI) to the adjusted (for baseline TUG score) model, whilst not reducing the intercept to zero, do not seem to improve the model. The deviance is unaffected and the PSI level variance varies only slightly. Considering less than 1% of TUG score variance between patients was attributable to PSIs once the model was adjusted for baseline TUG score, it is not surprising that the analysis cannot identify a candidate explanatory PSI characteristic.

However, the independent addition of patient-rated intensity and patient enjoyment, although lowering the PSI level intercept variance to zero, result in significant changes in deviance, indicating that patient-rated intensity and patient enjoyment are important predictors of post-intervention TUG time (after adjusting for baseline TUG). The patient level residual variance is lowered to 4.03 and 3.98 from 4.18, by the introduction of patient-rated intensity and patient enjoyment, respectively. 2%  $((4.18 - 4.03)/9.72)$  of patient-level variance has been explained by patient-rated intensity and 2%  $((4.18 - 3.98)/9.72)$  of patient-level variance has been explained by patient enjoyment.

#### **5.6.4.2.2 Chair rise**

The -2LL statistic was substantially reduced (1347 to 1092) when the control variable (baseline score) was introduced, giving a significant reduction in Deviance of 255 (Table 5.19). This suggests that baseline score is an important predictor of chair rise performance following the intervention. 76%  $((13.46 - 3.18)/13.46)$  of the patient level

variance and 12% of the between-PSI variance (0.89 to 0.78) has been explained by the effect of introducing baseline score. The ICC shows that almost 20% of the total variance is at the higher, PSI level after controlling for baseline chair rise. This is the only model where the proportion of PSI level variance has increased following adjusting for baseline score. The model is further improved by the addition of the predictor variable patient-rated intensity. The -2LL is reduced from 823 to 811. The effect of patient-rated intensity is significant ( $p=0.003$ ). Specifically, those patients who stated that the exercise was 'far too easy' have higher chair rise scores after the exercise intervention versus the reference category 'a little too hard'. The residual variance has decreased from 5.23 to 4.88, suggesting that 3%  $((5.23-4.88)/13.19)$  of patient-level variance has been explained by patient-rated intensity.

The introduction of patients' enjoyment of classes has a similar effect to intensity. The change in deviance is significant and the associated ICC suggests that about 10% of the total variance between patients' chair rise performance is at the PSI level. The effect of patient enjoyment is significant ( $p=0.009$ ). As with intensity, there is no reduction in the intercept variance (in fact, it increases) suggesting that this predictor is not operating at the therapist level. Patient enjoyment increased the PSI-level variance which indicated that there was a correlation between the patient-level predictor variable that was added and PSI-level variance. The differences that patients showed from their PSI-group mean chair rise score was related to how much variance in chair rise there was in that PSI-group of patients. Variation in chair among patients grouped by PSI was hidden in the unconditional model due to the absence of the predictor (patient enjoyment) and was revealed following controlling for this patient-level predictor variable. Patient enjoyment was operating at the patient-level and not at the PSI level.

The independent addition of all of the other predictor variables to the adjusted model (except PSI experience which reduces the intercept to zero) whilst not reducing the intercept to zero, do not improve the model.

**Table 5.19, Multilevel modelling; chair rise**

Model	Deviance (-2LL), df	Change in deviance, df	Is change in deviance significant ?	Effect of variable p-value	Patient level residual variance	PSI level intercept variance	ICC
Unconditional	1347, 1	-	-	-	13.46	0.89	0.062
Unconditional (n=196)	1006, 1	-	-	-	13.19	0.37	0.016
Baseline chair rise	1092, 2	255, 1	Y	<0.001	3.18	0.78	0.197
Baseline chair rise (n=196)	823, 2	183, 1	Y	<0.001	5.23	0.48	0.084
Baseline chair rise, PSI age	1092, 3	0, 1	N	-	5.15	0.31	0.057
Baseline chair rise, PSI gender	1090, 3	2, 1	N	0.104	5.14	0.21	0.039
Baseline chair rise, patient-rated intensity	811, 5	12, 3	Y	0.003	4.88	0.53	0.098
Baseline chair rise, patient enjoyment	813, 5	10, 3	Y	0.009	5.02	0.60	0.107
Baseline chair rise, PSI quality	1089, 4	3, 2	N	0.112	5.17	0.14	0.026
Baseline chair rise, PSI attendance	1092, 3	0, 1	N	-	5.15	0.31	0.057
Baseline chair rise, dose	1090, 3	2, 1	N	0.104	5.07	0.37	0.068
Baseline chair rise, PSI experience	1083, 3	9, 1	Y	0.001	5.14	0*	-

0\* = therapist level variation indistinguishable from patient level variation

#### 5.6.4.2.3 Romberg

The -2LL statistic was substantially reduced (1496 to 1421) when the control variable (baseline score) was introduced, giving a significant reduction in deviance of 75 (Table 5.20). This suggests that baseline balance score is an important predictor of balance performance following the intervention. 27%  $((23.89-17.45)/23.96)$  of the residual variance has been explained by the effect of adding baseline score. The model is

further improved by the addition of patient-rated intensity. The -2LL is reduced from 1072 to 1062. The effect of patient-rated intensity is significant ( $p=0.009$ ). Specifically, those patients who stated that the exercise was 'far too easy' and 'a little too easy' have higher balance scores after the exercise intervention when compared with the reference category 'a little too hard'. The patient-level residual variance was reduced by the introduction of the 'intensity' predictor variable, explaining 5% of variance at this level. PSI-level variance was unaffected by the addition of 'intensity' suggesting that there is no therapist effect for this variable.

The model is also improved by the addition of the enjoyment predictor variable. The -2LL is reduced from 1072 to 1062. The effect of patient enjoyment is also significant ( $p=0.009$ ). Patient enjoyment increased the PSI-level variance which indicated that there was a correlation between the patient-level predictor variable that was added and PSI-level variance. Patient enjoyment was operating at the patient-level and not at the PSI level.

The independent addition of all of the other predictor variables to the adjusted (for baseline score) model, whilst not reducing the intercept to zero, do not improve the model. The deviance is unaffected and the PSI level variance varies only slightly. These variables are not important predictors of post-intervention Romberg score.

**Table 5.20**, Multilevel modelling; Romberg

Model	Deviance (-2LL), df	Change in deviance, df	Is change in deviance significant?	Effect of variable p-value	Patient level residual variance	PSI level intercept variance	ICC
Unconditional	1496, 1	-	-	-	23.96	0.54	0.022
Unconditional (n=196)	1125, 1	-	-	-	24.62	0.18	0.007
Baseline Romberg	1421, 2	75, 1	Y	<0.001	17.54	0.66	0.036
Baseline Romberg (n=196)	1072, 2	53, 1	Y	<0.001	18.51	0.12	0.006

Baseline Romberg, PSI age	1419, 3	2, 1	N	0.104	17.44	0.60	0.033
Baseline Romberg, PSI gender	1421, 3	0, 1	N	-	17.56	0.61	0.034
Baseline Romberg, patient-rated intensity	1062, 5	10, 3	Y	0.009	17.54	0.13	0.007
Baseline Romberg, patient enjoyment	1062, 5	10, 3	Y	0.009	17.94	0.32	0.018
Baseline Romberg, PSI quality	1420, 4	1, 2	N	0.242	17.58	0.52	0.029
Baseline Romberg, PSI attendance	1412, 3	9, 1	N	0.001	17.03	0.43	0.025
Baseline Romberg, dose	1418, 3	3, 1	N	0.051	17.43	0.54	0.030
baseline romberg, PSI experience	1421, 3	0, 1	N	-	17.54	0.66	0.036

#### 5.6.4.2.4 Falls rate during the intervention

Falls rate has no equivalent baseline score and therefore cannot be adjusted like the functional assessments. The -2LL statistic was reduced (838 to 824) when the predictor variable patient enjoyment was introduced, giving a significant reduction in deviance ( $p=0.001$ ) (Table 5.21). This suggests that patient enjoyment of the intervention was an important predictor of falls rate during the intervention. 59% of the PSI level variance has been explained by the effect of adding patient enjoyment (see Table 5.22). This variable had a small effect at the patient level, explaining 4% of the patient level variance (see Table 5.22).

The independent addition of all other predictor variables to the model, whilst not reducing the intercept to zero, did not improve the model. The deviance was

unaffected indicating that these variables were not found to be important predictors of falls rate.

**Table 5.21, Multilevel modelling; falls rate during the intervention**

Model	Deviance (-2LL), df	Change in deviance, df	Is change in deviance significant ?	Effect of variable p-value	Patient level residual variance	PSI level intercept variance	ICC
Unconditional	1318, 1	-	-	-	7.35	0.11	0.015
Unconditional (n=196)	838, 1	-	-	-	4.26	0.22	0.049
PSI age	1317, 2	1, 1	N	0.242	7.33	0.10	0.013
PSI gender	1318, 2	0, 1	N	-	7.35	0.10	0.013
Patient-rated intensity	838, 4	0, 3	N	-	4.25	0.22	0.049
Patient enjoyment	824, 4	14, 3	Y	0.001	4.11	0.09	0.021
PSI quality	1315, 3	3, 2	N	0.112	7.35	0.02	0.003
PSI attendance	1315, 2	3, 1	N	0.051	7.33	0.04	0.005
Dose (patient attendance)	1318, 2	0, 1	N	-	7.35	0.10	0.013
PSI clinical experience	1317, 2	1, 1	N	0.242	7.38	0.04	0.005

#### 5.6.4.3 Summary; multilevel modelling

For falls rate during the intervention, Romberg balance score, and timed chair rise unconditional models, the intra-class correlation (ICC) was low indicating that there was more variation between patients within therapist groups than in mean outcomes seen for different therapists (Table 5.22). The Timed-up-and-Go (TUG) unconditional model ICC, was very high (Table 5.22), however, further investigation revealed that this difference was due to an outlier (Figure 5.11). When the functional assessment models (Romberg balance, timed chair rise and TUG) were adjusted for baseline score to control for clustering by practice, baseline score was found to be a statistically significant predictor of post-intervention score. Most of the variance was explained at



the patient level and little was explained by differences between the therapists (Table 5.22). The only significant improvements to the adjusted functional assessment models (Tables 5.16, 5.19 and 5.20) were seen when patient-rated intensity and patient enjoyment of the exercise sessions were introduced, therefore, only these two variables have been shown in Table 5.22. The p-values for both variables were significant for all functional assessment models, however, the proportion of variance at the PSI level explained by the addition of the explanatory variable (intensity or enjoyment) was zero. Patient-rated intensity and patient enjoyment of the exercise sessions were operating at the patient level only. In contrast to this, the proportion of therapist level (intercept) variance explained by fitting enjoyment to the falls rate model was 59%. However, there was only 5% of variance at the PSI level in the unconditional model to be explained. This small therapist effect is most likely a chance finding due to multiple testing. This will be discussed further in Chapter 6, section 6.2.3.2.

Dose was investigated in each multilevel model described in sections 5.6.4.2.1 to 5.6.4.2.4. The addition of dose to the model in each case resulted in no significant change to the deviance, suggesting that the dose of the exercise intervention was not a predictor of any of the falls outcomes; falls rate, and changes in falls risk factors, as indicated by the functional assessments.

**Table 5.22, Summary of multilevel modelling analyses**

		<b>Falls rate during the intervention</b>	<b>TUG</b>	<b>Chair rise</b>	<b>Romberg</b>
Unconditional model	ICC - proportion of variance at PSI level	0.049 (5%)	0.81 (81%)	0.062 (6%)	0.022 (2%)
Adjusted model	ICC - proportion of variance at PSI level	-	0.003 (<1%)	0.197 (20%)	0.036 (4%)
	Proportion of patient level (residual) variance explained	-	((10.52-4.20)/10.52) 60%	((13.46-3.18)/13.46) 76%	((23.96-17.54)/23.96) 27%
	Proportion of therapist level (intercept) variance explained	-	((43.47-0.01)/43.47) 99.9%	((0.89-0.78)/0.89) 12%	((0.54-0.66)/0.54) Zero
	Significance	-	p=<0.001	p=<0.001	p=<0.001
Patient-rated intensity	ICC - proportion of variance at PSI level	No significant change in deviance	0*	0.098 (10%)	0.007 (<1%)
	Proportion of patient level (residual) variance explained		((4.18-4.03)/9.72) 2%	((5.23-4.88)/13.19) 3%	((18.51-17.54)/24.62) 4%
	Proportion of therapist level (intercept) variance explained		0*	((0.48-0.53)/0.37) Zero	((0.12-0.13)/0.18) zero
	Significance		p=0.049	p=0.003	p=0.009
Enjoyment	ICC - proportion of variance at PSI level	0.021 (2%)	0*	0.107 (11%)	0.018 (2%)
	Proportion of patient level (residual) variance explained	((4.26-4.11)/4.26) 4%	((4.18-3.98)/9.72) 2%	((5.23-5.02)/13.19) 2%	((18.51-17.94)/24.62) 2%
	Proportion of therapist level (intercept) variance explained	((0.22-0.09)/0.22) 59%	0*	((0.48-0.60)/0.37) Zero	((0.12-0.32)/0.18) zero
	Significance	p=0.001	p=0.003	p=0.009	p=0.009

0\* = therapist level variation indistinguishable from patient level variation

## **5.7 Chapter conclusions**

For this population it was not possible to demonstrate an effect for dose nor any therapist characteristic in relation to either falls rate or falls risk factors. The main trial was not set up to investigate therapist effects and therefore was not powered for this. However, some of the effect sizes in the unconditional models show that there was some variance between patients grouped by PSIs. Effect sizes in standardised exercise interventions (such as FaME), especially within the research setting (as opposed to clinical practice), would not be expected to be large.

## **Chapter 6**

### **Discussion**

#### **6.1 Chapter summary**

This chapter begins with a summary of the findings from my therapist and dose effects study. I go on to discuss the meanings and implications of these results in relation to other research and clinical provision of falls prevention exercise. The strengths and limitations of my study are discussed, as well as the implications for further research.

#### **6.2 Summary of findings**

##### **6.2.1 Falls prevention**

The FaME intervention was effective at reducing falls in older people recruited through general practices in the ProAct65+ study in the year after cessation of the intervention. As well as this, significantly fewer injuries caused by falls occurred in the FaME group compared with usual care during the 24-week intervention period and in the combined intervention and first year post-intervention period. It appears that although the total number of falls (injurious and non-injurious) during the intervention was not significantly different between groups, the severity of the falls (in terms of the number of injuries sustained) was lower in the FaME group, suggesting that, although still falling in the intervention period, the FaME subjects were less likely to injure themselves. This might have been a transitional stage towards falling less often, which occurred in the FaME group in the first post-intervention year. Functional outcomes (TUG, timed chair rise, functional reach) did not improve despite a decrease in falls (Chapter 4, Section 4.10). Confidence in Maintaining Balance (CONFbal) was significantly improved at 12 months post-intervention in the FaME group compared with the control group.

There was a non-significant reduction in the falls incidence rate in the OEP arm compared with usual care in the 12 months following the cessation of the intervention. OEP exercise appears less effective at reducing falls in this functionally more able population of older adults. Functional outcomes did not improve, however, CONFbal

was significantly improved at 12 months post-intervention in the OEP group compared with the control group.

There was a reduction in injurious falls in the usual care group over the whole 2.5 years of the study and this reduction appears to be greater than the decline in injurious falls in either intervention arm. This may be a true finding, however, there are two alternative explanations for this. Firstly, response bias, as a higher percentage of falls diaries were not returned in the usual care arm (41%) compared to the OEP (35%) or FaME (37%) arms. We previously reported that participants at higher risk of falls were less likely to return falls diaries than those at lower risk (Perry et al., 2012). Secondly, attrition bias, as a slightly larger percentage of the usual care arm (31%) dropped out from the trial compared to the OEP (29%) or the FaME (27%) arms. We previously reported that those who dropped out were at higher risk of falls in terms of number of co-morbidities, number of medications and poorer performance on a range of functional measures (Iliffe et al., 2014).

### **6.2.2 Therapist effect study; bivariate analyses**

The baseline falls rate was similar in both the therapist and dose effect (TDE) group and those in the FaME arm who were excluded from the therapist effect analysis due to having no contact with a therapist throughout the trial duration (non-TDE group) (Chapter 5, Table 5.2). This makes the therapist effect analyses results potentially more generalisable.

Bivariate analyses conducted on the TDE group suggested that patients' enjoyment of classes was strongly associated with reduced falls at all time-points. Those who thought the classes were enjoyable had a significantly lower falls rate both during and after the intervention compared to those who were ambivalent about the classes (Chapter 5, section 5.6.2.3.1). Multi-level modelling was used to explore whether patients' enjoyment of classes differed between therapists (see section 6.2.3.2). There was also a significant difference in functional assessment scores between groups of patients according to their rating of exercise class intensity. Those who thought the exercise was too easy had better balance, leg power and gait. These are not surprising

findings. However, the rationale for including patient-rated exercise intensity was to explore possible therapist effects, which would not be achieved in the bivariate analysis. Multi-level modelling was used to provide further explanation. Furthermore, the multi-level modelling analyses were adjusted for baseline functional assessment scores to rule out the possibility of allocation bias.

We found a significant negative association between regular attendance of PSIs to their allocated sessions and post-intervention Romberg balance score (higher balance scores, indicating superior balance, were associated with poorer PSI attendance) (Chapter 5, Table 5.9), and chair rise group means were found to be statistically significantly associated with the timing of the instructor's PSI training (patients whose PSI was trained as part of the trial had a significantly higher mean number of chair stands compared to those whose PSI was trained before the trial) (Chapter 5, Table 5.11). We believe these are most likely chance findings due to multiple testing because neither predictor was significantly associated with any other patient outcome. Moreover, bivariate analyses were not adjusted for baseline functional assessment score therefore findings are preliminary and exploratory. The multilevel modelling analyses findings are the definitive results of my therapist and dose effects study.

### **6.2.3 Therapist effect study; multilevel modelling analyses**

In the population studied (community-dwelling, over 65-year-olds) it was not possible to demonstrate a consistent therapist effect in relation to either falls rate or falls risk factors for any therapist characteristic. It was also not possible to demonstrate a dose effect in relation to either falls rate or falls risk factor outcomes. Attendance varied at the individual level, and in many cases was poor, but did not vary between therapist groups.

#### **6.2.3.1 Multilevel modelling; unconditional models**

There was no evidence of any variation between therapists for three patient outcomes; 1) falls rate in the 12 months following the intervention, 2) falls rate in the combined intervention and 12 months following the intervention period and 3) functional reach. For three further patient outcomes; 1) falls rate during the

intervention, 2) Romberg balance score and 3) timed chair rise, some variance was identified, but most of this was due to differences between patients, rather than differences between groups of patients assigned to different therapists. Differences were greater within groups rather than between groups. The Intra-class correlation (ICC) was low in all three models indicating that there was more variation between patients within therapist groups than in mean outcomes seen for different therapists. Although there was some between therapist variation it was small, statistically not significant and unlikely to be clinically important. The Timed-up-and-Go ICC, however, was very high indicating that patients grouped by PSI correlated strongly. Further investigation revealed that this difference was due to an outlier (Chapter 5, Figure 5.11).

Overall, the unconditional models lacked robust evidence of a therapist effect. However, some variance was not completely explained by differences between individuals. In addition, given that the trial attempted to standardise the delivery of the exercise intervention, large variation between therapists would be unlikely. These findings provided a rationale for the addition of predictor variables to the models in order to explore the reasons for the variance.

#### **6.2.3.2 Multilevel modelling; adding predictors to the models**

Baseline functional assessment scores were added to the models to control for clustering by practice. Baseline score was found to be a statistically significant predictor of post-intervention score for all functional assessments. Most of the variance was explained at the patient level and little was explained by differences between the therapists (PSIs) suggesting that baseline functional assessment scores were not markedly clustered by PSI. That said, there was still some variation between therapists requiring further exploration. For some outcomes the percentage of variance at the PSI-level had increased after adjusting for baseline score. Increases in level 2 variance resulting from the addition of level 1 predictors to the model suggests that there was a correlation between the variables. For TUG, the addition of baseline score to the model accounted for all variance at the PSI level, therefore leaving no

therapist effect to be explained. For falls rate there was no equivalent baseline score to introduce, therefore the unconditional model was used as the comparator.

Predictor variables were added independently to the adjusted models in order to explore therapist characteristics that explained the small remaining PSI-level variance. The only significant improvements to the model (Chapter 5, Tables 5.16, 5.19, 5.20 and 5.21) were seen when patient-rated intensity and patient enjoyment of the exercise sessions were introduced. Patient-rated intensity and patient enjoyment of the exercise sessions were significant predictors of change in functional assessment scores. However, both patient-rated intensity and patient enjoyment of the exercise sessions were mainly operating at the patient level. There was no evidence of any therapist effect associated with patient-rated intensity and patient enjoyment of the exercise sessions on falls risk factors.

Patient enjoyment of the exercise sessions was a significant predictor of falls rate during the intervention and for this outcome patient enjoyment of the exercise sessions was mainly operating at the PSI level; explaining 59% of the PSI level variance in falls rate and only 4% of the patient level variance. However, considering the amount of variance at the PSI level to be explained was only 5% of the total variance in the unconditional model, this appears to be, at best, a small therapist effect. However, we believe this may be a chance finding due to multiple testing, given that enjoyment was not grouped by therapist for any other dependent variable.

## **6.3 Relationship to other research**

### **6.3.1 Falls prevention**

The ProAct65+ study was the first evaluation of a short version (24 weeks) of the Falls Management Exercise programme (FaME) in the primary prevention of falls. FaME had previously been used in the secondary prevention of falls as a nine-month intervention (Skelton et al., 2005). The original FaME study (Skelton et al., 2005) showed a significant reduction in falls in community-dwelling, female, frequent fallers. The ProAct65+ population were community-dwelling over 65s, recruited through



general practice, who were not selected due to a history of falls. The reduction in the incidence of falls in the ProAct65+ population adds to the growing body of exercise intervention research (Korpelainen et al., 2006, Madureira et al., 2007, Means et al., 2005, Voukelatos et al., 2007, Sherrington et al., 2008, Sherrington et al., 2011) which indicates that falls prevention exercise programmes should be utilised to prevent first falls, rather than being viewed primarily as a treatment or rehabilitation option for those who have already sustained falls. As well as this, the finding regarding the effectiveness of the shorter FaME intervention in the general older population may be used as evidence for NHS exercise services to offer the shorter FaME programme if they are working with the general older adult population in the primary prevention of falls, rather than selected, frailer individuals who have a history of falls.

There is some evidence that increasing physical activity in those at high risk of falls can increase exposure to risk (Sherrington et al., 2011), so it is noteworthy that, in older people recruited through general practice (not frequent fallers), the FaME intervention did not increase the risk of falls alongside the increase in physical activity.

It is also of note that the functional outcomes did not improve despite a decrease in falls and an increase in physical activity. A psychological mechanism may account for the positive falls outcome, and this will be discussed later in this chapter (see section 6.3.2.3). Mean post-intervention improvement (compared with baseline score) for TUG, functional reach and tandem stance balance with eyes open and eyes closed was approximately 10, 10, 70 and 60 percent, respectively, in the original FaME trial (Skelton, data not published). As FaME incorporates the OEP balance exercises, it is relevant to note that the first OEP study reported a significant improvement in balance (measured by the 4-test balance score, which includes tandem stance assessment) in the exercise group compared with the control after 6 months of exercise (Campbell et al., 1997). The ProAct65+ population scored poorly in the functional tests at baseline in comparison to normal data (Chapter 4, Table 4.3) so there was room for improvement in both strength and balance. It is possible that other, more sensitive tests may have shown some improvement, such as more dynamic balance tests (Tinetti, 1986, Berg et al., 1995, Berg et al., 1992), compensatory stepping ability (Medell and Alexander,

2000), reaction time (Lajoie, 2004), or other components of fitness that were not tested. There was also no change in fear of falling as measured by Short FES-I, despite a Cochrane review suggesting that strength and balance exercise reduces fear of falling (Kendrick et al., 2014). However, only a very small percentage of people in the ProAct65+ population expressed concern about falling at recruitment.

### **6.3.2 Therapist effects**

#### **6.3.2.1 Standardised interventions**

Most contributions to the therapist effect literature are from the field of psychotherapy and report practitioner effects of between 1 and 12% (intraclass correlation coefficient; the proportion of variance resulting from the therapist), with a mean estimate of 8% (Kim et al., 2006). The literature review (Chapter 2) conducted as part of this study revealed a dearth of practitioner effects literature in the area of exercise. Previous studies of therapist effects in physiotherapy, however, suggest that therapist effects are less likely in treatments where emphasis is placed on the standardisation of treatment practice (Lewis et al., 2010). The ProAct65+ trial carefully monitored the fidelity of the FaME intervention via quality assurance (QA) site visits and required exercise instructors (PSIs) to action QA feedback to try to ensure that patients received a standard intervention, irrespective of the PSI delivering it. The lack of therapist effect identified in this study, may reflect the effectiveness of the standardisation procedures. Lewis and colleagues additionally comment that therapist effects are likely to be greater in interventions that are flexible. FaME is an inherently prescriptive intervention. This is in the most part due to the falls prevention exercise research field having already clearly identified an effective, and relatively inflexible, exercise prescription for reducing falls. There have been some studies that utilised the 'correct' type of exercise (balance and strength) which failed to show a reduction in falls, suggesting that the falls prevention exercise prescription is more complex than exercise type alone, and may involve other variables such as the length of the

intervention and the amount of progression in exercise intensity (Sherrington et al., 2008, 2011).

Another hypothesis is that therapist effects are more pronounced in interventions that allow more flexibility not only in the treatment selected but in the practitioner-patient interaction. Lewis (2010) states that interventions with a more psychosocial focus are likely to allow greater variability in patient outcome. The falls prevention exercise equivalent to the psychosocial interventions described to reduce back pain (Jellema et al., 2005, Hay et al., 2005) might be falls prevention exercise counselling or motivational interviewing as a pre-cursor to falls prevention exercise, and it would be interesting to study, if this approach was routinely used in clinical practice. In reality, it is likely that only the more self-motivated individuals access falls prevention exercise services, and those who are found to lack motivation are offered alternative, non-exercise interventions, such as home modifications.

Previous research has also suggested that conformity in practice is easier to achieve in smaller groups of therapists. Some exercise/physiotherapy therapist effect studies have studied therapist groups as large as 930 (Resnik and Hart, 2003), in comparison to only 12 PSIs in ProAct65+. The relatively small number of therapists may therefore have contributed to the failure to find therapist effects in this study; a possible type II error.

#### **6.3.2.2 Therapist characteristics studied**

My literature review included studies that have analysed the effect of various therapist characteristics including years of experience, advanced training, age and gender. Studies involving other 'therapists' (GPs, dentists, urologists, physicians) have also included attentive skills (Rasmussen and Andersen, 2005) and technical skills (Hemminiki, 1982). As our results showed some small therapist effects, but no explanatory factors, it may be that the characteristics studied did not include characteristics that make a difference to our patient outcome; falls rate reduction. Hawley-Hague (2014) suggests that exercise instructors' personality traits are

associated with participant adherence to group exercise classes. These less objective characteristics were not collected nor analysed in the ProAct65+ study.

Loss of adherence to the evidence-base ('therapist drift') is described as being an example of unplanned therapist variation (Waller, 2009). In theory, therapist drift was eliminated from the ProAct65+ trial by the QA procedures. However, the QA researchers observed that despite QA procedures there was a range of technical skill and strict adherence to the evidence base was varied, in particular in the progression of the programme. Some PSIs were confident about exposing participants to a higher level of risk than others and this did not appear to be an accurate reflection of group ability; 'cautious' PSIs progressed all of their groups across the course of the study to a similar, but lower, level than other PSIs. However, in the adapted FaME intervention, with this group of patients, the variation between therapists was insufficient to affect patient outcomes.

Ensuring that exercise is at the appropriate intensity for individuals within a group is suggestive of therapist skill. It is inherently more difficult to tailor for individuals in group exercise compared with tailoring in a one-to-one (personal training) exercise session. Ensuring exercise is enjoyable arguably encompasses many factors, including, but not exclusively, having rapport with individuals, tailoring exercises for individual needs and goals, being friendly and approachable, and seeking feedback from clients. One study found that client satisfaction with group exercise provision was found to be correlated with personal goals, productivity; such as the amount of exercise provided in the session, and group integration (Loughead and Carron, 2004). So enjoyment of group exercise may well be influenced by factors that would also have been grouped by therapists in my analysis, but that are not necessarily mediated by the therapist; for example, the rapport between group members or their liking of the exercise venue. All that can be concluded about enjoyment of exercise from my study is that enjoyment of a standardised falls prevention exercise intervention is a significant predictor of

falls-related patient outcomes, and that levels of enjoyment are not grouped by therapists.

#### **6.3.2.3 Functional assessments**

A possible conclusion to be drawn from the lack of change in objective balance and leg power measurements in the ProAct65+ FaME group compared with usual care, as well as by therapist, is that the functional assessments selected were not amenable to change in those whose baseline scores were already higher than, or at the higher end of the range, expected for older people with falls risk. This would fit well with the proven relationship between baseline physical activity or functional status and the benefit gained from increasing physical activity; those with a higher baseline physical activity or functional status benefit less from increasing their physical activity compared to those whose baseline status is low (Pate et al., 1995, Arem et al., 2015). Thus it may have been difficult to elicit measurable changes in balance/leg power in a 24-week falls prevention intervention in a general, older population (rather than a population of selected, frailer individuals with a falls history). Analysing therapist effects in future studies of falls prevention exercise should explore effects in those with a frequent falls history (as this will identify those with a lower baseline functional status) or alternatively in those with a high risk of future falls, identified, for example, by the Falls Risk Assessment Tool.

Given that there was no overall change in any of the functional assessments in the FaME group compared with the control, it is not surprising that there was no difference between therapists for these measures. The lack of change in balance and leg power as measured by the functional assessments, alongside the reduction in falls incidence in the FaME group compared with usual care, appears to suggest that the mechanism for reducing falls in the FaME group was not via improved balance and leg power (Chapter 4, section 4.10). The ProAct65+ results, however, did demonstrate improved confidence to maintain balance during some everyday tasks (measured by CONFbal) in the FaME group compared to usual care. This suggests that the mechanism for reducing falls in this population, and with a shorter intervention than the original FaME intervention, may have been as a result of improved psychological

factors pertaining to falls. The original FaME trial (Skelton et al., 2005) had shown a non-significant increase in falls during the intervention period for the exercise group compared with the control group, and this was thought to be due to the increased exposure to falls risk that the intervention itself posed. It may be that balance confidence had improved before the improvements in physiological falls risk factors such as balance and leg strength. Indeed, a systematic review of interventions for improving balance confidence reported that it can be significantly improved in as few as 5 weeks of exercise (Rand et al., 2011), whereas balance and strength improvements require in the region of 12-42 and 8-12 weeks of exercise, respectively (Dinan-Young and Skelton, 2012). Skelton and Dinan (1999) note that improvements in these risk factors can result in 'over confidence' which acknowledges the complex relationship between the physiological training effects of falls prevention exercise and associated psychological responses. What remains unclear is how the therapist's characteristics (including technical skills such as confidence to ambitiously, yet safely, progress exercises) impact on participants' confidence to maintain their balance. It may be that patients' confidence is improved simply by the sense of achievement gained from participation in (group) exercise that includes balance activities, irrespective of the level of challenge (intensity) of those exercises. This is potentially an interesting area for future research.

#### **6.3.2.4 Quality assurance of intervention**

One further discussion point is regarding the measure and definition of PSI 'quality' within this study. The trial quality assurance staff assessed the quality of the intervention delivered by each PSI according to pre-set national occupational standards that have been used historically by the fitness industry to formally assess trainee exercise instructors and thus award them a licence to practice. These criteria had already been modified and are used by Later Life Training ([www.laterlifetraining.co.uk](http://www.laterlifetraining.co.uk)) for assessing trainee PSIs in order to award them a licence to practice. The same criteria were further modified by the author (SG) for ProAct65+ to check the fidelity of the intervention and ensure standardisation. The tool was judged to be effective for this purpose and allowed QA staff to make recommendations to PSIs to improve fidelity and standardisation of the intervention.

This raises the question of whether the adapted assessment checklist was the most appropriate tool for assessing 'quality' of PSIs within the therapist effect study and that perhaps other methods could have been considered. For example, Hart and Dobrzykowski (2000) suggest that quality of treatment delivered by physical therapists can be measured by the amount of functional improvement per episode of care or per visit. This retrospective assignment of quality rating according to patient outcome, whilst being more objective, would not have worked for the therapist effect study, which was by design investigating the effects of therapists' characteristics on patient outcome, rather than assign the 'characteristics' once the treatment had concluded. However, further consideration might be given to how technical skills or therapists' 'quality' are judged. It seems important from the therapist effect outcomes relating to patient evaluation to at least ensure that the 'quality' of the PSI is evaluated not only by supervisors/line managers within the workplace, but also by the patients themselves. This is by no means a novel concept; the Royal College of Physicians included older people's satisfaction of falls services exercise provision within their 2012 audit (Buttery et al., 2013), but my study suggests that patients may be able to give valuable insight into the likely effectiveness of the exercises to reduce falls.

#### **6.3.2.5 Clinical experience**

From the broader perspective of any exercise research, it is typical for therapists/instructors with the most clinical experience to be invited to participate (Whitman et al., 2004). An assumption is made that greater experience leads to better patient outcomes. My study suggests that previous experience of delivering the FaME exercise programme need not be a pre-requisite for PSIs' participation in exercise research, and that provided research therapists/instructors are qualified, trial-trained and engage in self-evaluation and modification of their practice via quality assurance observations, there may not be therapist-level variations in patient outcome. This may be due to the fact that the FaME intervention is standardised, with pre-set progressions, and is not as reliant on clinical decision-making/clinical reasoning as other interventions such as physiotherapy. Whitman and colleagues suggest that less experienced therapists may indeed be more capable than experienced therapists at learning and administering standardised techniques (Whitman et al., 2004).

Inexperienced exercise instructors, may, however, find it challenging to identify which patients are most likely to benefit from the intervention. In research this would be controlled by trial inclusion and exclusion criteria, however, in a falls exercise service, the detection and assessment of those at high risk of falls should possibly be the role of more experienced (and PSI trained) exercise instructors or physiotherapists. The Later Life Training PSI manual highlights the differing roles of the clinical exercise specialist and the physiotherapist within a falls rehabilitation multidisciplinary team, and notes that the physiotherapist is responsible for the physical assessment of the patient following a fall (Dinan-Young and Skelton, 2012).

#### **6.3.2.6 Time-points studied**

Hawley-Hague and colleagues (2014) investigated associations between group exercise instructors' characteristics and attendance and adherence. They showed that different instructor characteristics were important at different time-points, with male gender, younger age and a greater number of years of teaching experience being associated with better adherence in the first 3 months. Across 6 months, however, only one personality trait ('conscientiousness') was associated with greater adherence. When the primary outcome is falls incidence (rather than attendance), there does not seem to be a logical rationale for investigating therapist characteristics within shorter blocks of time (say, 3-monthly) across the entire intervention, given that the reduction in falls incidence was only evident in the 12 months following the cessation of the exercise classes. Falls reduction, unlike attendance, is not achieved immediately as it requires physiological change resulting from the exercise (the 'training effect'). However, as seen in the Hawley-Hague (2014) analysis, therapist effects that were significant in only a proportion of the whole intervention, and are hidden ('diluted' and therefore or no longer significant) when a longer period is investigated. In future therapist effects in falls prevention exercise studies it might be of interest to see if therapist skills, especially relating to selection of exercise intensity and rate of progression, are associated with increased exposure to risk and therefore increased falls in the first 3 months of the intervention. There was a small, non-significant increase in falls in the intervention group in the early intervention phase of the original FaME trial (Skelton et



al., 2005). Investigating exposure to risk in this way could inform exposure to risk in clinical practice.

#### **6.3.2.7 Gender**

The gender and age of the FaME therapist appeared to have no effect on the patient outcomes that were investigated in the ProAct65+ trial. Existing therapist effect literature in the field of psychotherapy suggests that neither gender nor age affect patient outcomes (Anderson et al., 2009, Cella et al., 2011), so this finding is not surprising.

#### **6.3.3 Adherence & dose**

The 2005 FaME trial reported that 79% of subjects attended more than 75% of classes (Skelton et al., 2005). By comparison, adherence was poorer in the ProAct65+ FaME group with only 40% (n=150) attending 75% or more of classes. The original trial (Skelton et al., 2005) recruited frequent fallers who may have been more motivated to attend falls prevention exercise sessions, whereas ProAct65+ excluded them.

In my dose analysis, a total volume of supervised exercise achieved by each patient was well calculated from attendance registers (see Chapter 3, section 3.4.5.5). This was referred to as the 'dose' of exercise (per patient) in my study. The rationale for this has been described earlier in the thesis (Chapter 1, section 1.12). I found that dose was neither an independent predictor of falls rate and falls risk factors nor was it mediated by the therapist. Some therapist factors have been shown to positively influence attendance/adherence to exercise (Seguin et al., 2010, Hawley-Hague et al., 2014). In addition, the duration of the exercise intervention in hours is thought to have an optimal dose of more than 50 hours for falls reduction (Sherrington et al., 2008, Sherrington et al., 2011). In my study it was hypothesised that therapist factors would influence attendance thus affecting dose, which in turn would have an influence on falls outcome. Attendance was varied at the individual level, and in many cases was poor, but was not grouped at the therapist level. This finding is not in keeping with previous research showing that exercise leaders can influence adherence, and it can be concluded that in this standardised falls prevention exercise programme, attendance

(and therefore, dose achieved) by community-dwelling over 65s may not be mediated by therapist factors.

However, Sherrington (2011) stated that there was not a clear cut-off point for dose in falls prevention exercise trials, suggesting that some trials with shorter interventions may have been effective. I looked for trials with interventions of less than 50 hours in lower risk participants that were included in either the 2008 or the 2011 Sherrington meta-analysis and found four that reported a reduction in falls (or falls risk factors) in the intervention group compared with the control (see Chapter 1, section 1.12). These studies suggest that in some low risk populations, falls can be reduced with between 16 and 40 hours of exercise. Future studies investigating the dose of falls prevention exercise in the general older population might consider categorising the total dose of exercise achieved by each patient at a range of cut-off points (40 hours, 30 hours, 20 hours, for example). This would be in keeping with the Sherrington meta-analyses.

#### **6.4 Limitations and strengths**

As this PhD was nested within an RCT that was designed for another purpose there are a number of inherent limitations in the study design that require caution in the interpretation of the results. My study was not powered to detect therapist effects. Power calculations were based on the primary endpoint of the main ProAct65+ trial. The power calculation I presented in Chapter 5, section 5.5 suggested that my study was underpowered. I found no significant difference between therapists, but it is possible that there were insufficient data to detect any differences, especially given that the differences were likely to be small anyway, in a standardised, quality assured intervention. Underpowered trials can also cause Type I errors (when the null hypothesis is incorrectly rejected), but this is not the case here, as no effect size was detected.

The cluster-randomised design of ProAct65+, with all patients at each group practice being randomised to the same study arm, meant that there was the potential for allocation bias (a difference between the patients in the intervention and control

groups). Patients from a group practice typically all live within the same geographical location and factors such as the deprivation level within that location or the density of population (urban versus rural), could have had an impact on health and also potentially on exercise habits, physical function and physical fitness. The design of ProAct65+ may therefore have resulted in groups of individuals with poor health and physical function being randomised into the same study group. In the main trial analyses this was controlled for by adjusting for practice deprivation and practice size. There was the same potential for bias within patients grouped by therapist, given that each FaME class contained a group of patients from the same GP practice, and that the same PSI would often teach more than one group of patients from that practice. The latter occurred for pragmatic reasons and was simply the result of the most cost-effective way of deploying the trial PSI workforce; less travel expenses were incurred if the PSI taught two consecutive sessions in the same community venue on the same day, plus back-to-back classes are more attractive to the PSI as they reduced travel time, for which PSIs were not paid. The therapist effect (TE) analyses should therefore have also controlled for this potential allocation bias. In all the falls risk factor analyses this was achieved by introducing baseline functional assessment scores (for TUG, functional reach, timed chair rise and Romberg balance) into the multi-level modelling. For the falls rate analyses this was not possible as there is no equivalent, accurate baseline measure to add to the model. This is a weakness of the TE analysis, and should the results have demonstrated a significant difference in falls at the therapist level, this result would need to have been interpreted with caution, or otherwise, practice deprivation and practice size would need to have been controlled for. The exercise instructors recruited to the study were not selected on the basis of their characteristics therefore there was not a balance of these characteristics in the sample of instructors. Only 2 (17%) instructors, for example, were male, and only 2 (17%) were physiotherapists. Additionally, the number of patients allocated to each instructor was varied (ranging from 8 to 71 patients). These factors, along with the small sample of therapists, would have limited the generalisability of the results should a therapist

effect have been shown. In a primary study of therapist effects this would need to be considered at the design stage.

The therapist effect analysis used falls incidence as the primary endpoint, thus was reliant on the quality of the falls diary data. We found that the falls diary used in ProAct65+ was frequently incorrectly completed, or not returned (Perry et al., 2012), resulting in over a third of missing diary data. Missing data was accounted for by calculating a time at risk of falls for each patient based on the number of diaries they completed. This method assumes the rate of falls would be constant over time for each patient. This is unlikely, but was considered by the ProAct65+ team better than assuming that those who did not return falls data did not fall. The potential impact of measurement bias, when data are missing at random, is to underestimate effects. When data are missing not at random, as we found in the ProAct65+ trial (see Chapter 4, section 4.9), underestimation is not necessarily the case. In a primary study of therapist effects, therefore, methods for minimising missing falls data would need to be considered.

As the multilevel modelling results showed some evidence of possible therapist effects, but as the therapist characteristics studied did not seem to explain these weak effects, it may be that the characteristics studied did not include those that make a difference to the patient outcome; falls rate reduction. Hawley-Hague (2014) suggests that exercise instructors' personality traits are associated with participant adherence to group exercise classes. Less objective therapist characteristics, such as personality traits or communication skills, were not collected nor analysed in this study. Particularly in relation to the dose effect investigation, this is a limitation of this study. It also may have helped to make sense of the effect of patients' enjoyment of the sessions.

A final limitation concerns the method of calculating a therapist 'quality' categorisation from quality control data. The methodology used has face validity but was not subject to other validity testing or formal reliability testing, although I did less formally evaluate it by having two independent raters. It would have been more defensible if I

had reported on the level of rater agreement and calculated a Kappa statistic, or at least reported the percentage where raters allocated the PSI to the same 'quality' categorisation group. However, although this would have been more robust, it is not possible to do this retrospectively as the two raters met to agree a consensus rating, therefore the only rating available is the combined, agreed rating for each PSI.

As this PhD was nested within a large RCT, it shares some of the strengths of the main study (Iliffe 2014). It was one of the largest samples achieved in a therapeutic exercise trial to date, as well as recruiting the general older population which allows broader generalisability of any findings than if a specific patient population (such as those with diabetes) had been selected. The London site in particular recruited a diverse older population in terms of ethnicity and first language. The ProAct65+ trial was also multi-centre thus achieving a broader geographic spread of participants than would have resulted from a single site, further improving generalisability of findings.

The mean age in the therapist effect group was 73 years, with 16% aged 80 years or over. 63% were female and 11% were non-white. For more detail regarding the baseline characteristics of the therapist effect group see Chapter 5, section 5.3. The therapist effect group demographic was very similar to the entire ProAct65+ population (Chapter 4, section 4.3). My therapist effect sample reflected the general population of older people living in England in terms of age and ethnicity (Health & Social Care Information Centre, 2014). Although men were slightly under-represented, this should not limit the inferences that can be made. My sample for the TE analysis were on average more functionally able than the FaME subjects who were excluded from the TE group. However, the ProAct65+ population was found to score more poorly than published average values for the functional assessments, thus making my TE group more representative of the general population of older people in terms of physical function. In terms of falls rate, which was my primary outcome, the TE group matched the larger ProAct65+ sample. That said, frequent fallers were excluded from

ProAct65+ therefore any inferences from the findings cannot be extrapolated to this population.

Another strength is that my exploratory work on therapist effects is completely novel; it is the first time the effect of the therapist has been studied within falls prevention exercise and indeed is the first time therapist effect has been investigated within the exercise trial setting.

## **6.5 Implications for service provision**

There are a number of implications from the findings from my PhD for service provision and practice.

### **6.5.1 Clinical practice versus research**

In my thesis I have shown that it was not possible to demonstrate a consistent therapist effect on falls outcomes for any therapist characteristic measured in the ProAct65+ trial. This implies that when an exercise intervention is delivered in a very standardised way in the strict research setting, minimal differences between therapists may be observed. In clinical practice, however, PSIs often work alone, without assistants or co-workers, when delivering the FaME intervention in the community. It is possible that the intervention could be diluted due to poor adherence to the evidence based exercises, and that a greater variability in how PSIs deliver the intervention occurs. This variation may occur for several reasons, including feedback from participants (for example, they express either enjoying or not enjoying particular exercises, which influences the PSI's exercise selection), PSI boredom (due to repeating the same exercises) and lack of continuing professional development and peer/expert support. The National Quality Assurance Framework 2001 provides guidelines for exercise referral (ER) systems and specifically includes falls as part of the remit of ER schemes and therefore the scope of the guideline (Exercise Referral Systems: A National Quality Assurance Framework, 2001). Under the umbrella of evaluation there are some specific recommendations regarding the audit of clinical exercise instructors delivering the interventions. Continuing Professional Development (CPD) is advised

(and indeed is a requirement of the Register of Exercise Professionals (REPs)) plus there is a recommendation that ER services should adopt “cycles of peer or ‘expert’ observation, discussion and decision to modify practice, change in practice, and reflection of the change”. It is considered good practice for there to be annual or bi-annual quality assurance observations of intervention delivery and indeed many NHS exercise services follow this model.

The tool used to ensure FaME intervention fidelity in this study was developed from the NVQ Level 4 assessment documentation used by Later Life Training in the awarding of the PSI qualification. This in turn is based on Level 2 (entry level) assessment documentation for all health-related exercise qualifications and contains assessment criteria (against which candidates are judged) that were agreed by consensus recommendation from experts in health-related exercise (Appendix). The assessment criteria are vocational and as such are believed to reflect the practical exercise delivery skills required to achieve success in the exercise industry. There can be a mis-match between what industry experts and end-users consider to be desirable qualities or skills needed to do the job to a high standard. For example, exercise session participants might rate exercise instructors on their approachability or enthusiasm, whereas exercise and fitness managers may favour correct exercise technique advice. When exercise is delivered to mainstream, healthy populations, adherence to evidence is rarely considered a priority by instructors, participants or managers. However, in higher risk patient-populations, with specific exercise outcomes, such as falls prevention, understanding evidence and applying it to practice is more important and more likely to be assessed (Exercise Referral Systems: A National Quality Assurance Framework, 2001).

In this study, as judged by the quality assurance documentation, there appeared to be a noticeable range in ‘quality’ of intervention delivery. This was most evident between those PSIs rated as ‘high quality’ in comparison to those rated as ‘low quality’. The QA checklists identified that those PSIs with lower quality delivery skills were less able to 1) adhere to evidence-based exercises and evidence-based exercise technique, 2)

progress exercises sufficiently for the participant group, 3) provide appropriate teaching points to ensure safe and effective exercise performance and 4) maintain adequate research paperwork; for example, class registers. However, the quality of delivery as judged by the QA researchers using the QA documentation did not appear to have a significant effect on falls rate nor on any of the risk factors for falls, such as balance. There are several potential conclusions that can be drawn from this. Firstly, the trial attempted to standardise delivery of the intervention across sites and between PSIs by providing trial 'top-up' training (which included some practical revision of evidence-based exercises and associated technique) and by making QA site visits to each PSI and providing feedback and action points for any observed standardisation issues. If these measures were successful, the difference between PSIs would be minimised, suggesting that the absence of a therapist effect was as a result of standardisation procedures that were highly effective. However, if this were the case, it seems unlikely that QA staff would have reported noticeable differences in PSI quality. The second possible conclusion, therefore, is that the factors (or skills) being assessed by the QA documentation are not, in fact, important factors associated with the successful delivery of the intervention. This seems counterintuitive given that the QA documentation was written and used by 'experts' who are knowledgeable about the evidence. For example, the Sherrington (2011) meta-analysis concludes that, amongst other things, falls prevention exercise should include highly challenging balance exercises. PSIs who were cautious in their selection of balance exercises were therefore perceived by the QA staff to be delivering balance exercises that were not sufficiently challenging and this would have been recorded under criterion number 15 on the PSI QA Checklist (see Chapter 3, Figure 3.3). Discussion with PSIs following the QA site visits raised the question of whether 'highly challenging' was understood as the level of the exercise in its own right, or in relation to the functional ability of the individual performing it. For example, dynamic (moving) balance exercises that also reduce the client's base of support, such as walking on the toes, would most likely be categorised as high level balance exercises irrespective of client ability. A more functional activity, such as a sit to stand transfer, might provide a balance challenge for frailer, more unsteady individuals, but would not be considered a highly challenging balance-retraining task in its own right. Sherrington (2011) states that balance



exercises which provide a moderate to high level of challenge 1) reduce the base of support, 2) reduce the upper limb support and 3) change the centre of gravity. This appears to reinforce the view that although any exercise should be tailored to the individual, balance exercises that are effective must meet certain challenge criteria, and that cautious PSIs may not be finding the right compromise between safety and exposure to risk to achieve this effectiveness. Despite this theoretical defence for the 'quality' assessment of PSIs, PSI quality was not associated with improved falls outcomes in this study.

Furthermore, despite the QA finding that PSIs differed in 'quality' and that some PSIs were cautious exercise progressers, there were no differences in patients' perception of the intensity of the exercises between PSIs. Patients in this trial may have been able to effectively judge exercise intensity because they were not the frailer, older population typically referred to falls services and as such were more cognitively intact. Also those with cognitive impairment that would interfere with group exercise class participation were excluded. Provided, therefore, that the PSI attempts to give an indication of how hard the exercise should feel, and offers a range of intensity options, there is every reason to assume that the best person to make the decision about which option to select, is the patient themselves. This is in line with theory around rate of perceived exertion (RPE) use which suggests that, with education, people are very good at judging how hard they are working (Borg, 1982). This method of monitoring intensity is recommended, in preference to other methods, such as tracking heart rate, when exercising older adults, as it is not confounded by disease and medication. PSIs should continue to seek feedback from participants with regard to intensity and effort levels.

### **6.5.2 Quality assurance in clinical practice**

Another discussion point about the quality of the FaME delivery is the issue of service evaluation. As it is considered good practice for there to be quality assurance of intervention delivery (Exercise Referral Systems: A National Quality Assurance Framework, 2001), and many NHS exercise services follow this model, these findings may have implications for how this is implemented. One conclusion might be that the

training and QA model used in the ProAct65+ study, and therefore also in my therapist effect study, is effective in sufficiently standardising PSIs to achieve primary prevention of falls outcomes in the general older adult population, although further investigation would be needed to provide more conclusive evidence.

My dose effect study looked at the effect of different doses of FaME exercise on falls outcome and found no association. Neither lower nor higher doses were found to positively or negatively affect falls incidence rate. However, the intention-to-treat analysis showed a reduction in falls incidence in the exercise group compared to the control group, in the 12 months following the end of the intervention. Furthermore, meta-analyses of falls exercise research suggest that a minimum dose of 50 hours of exercise is required for falls reduction patient outcomes (Sherrington et al., 2011). ProAct65+ utilised a shorter intervention (24 weeks versus the original 9 months) and adherence was poorer than expected (40% achieving 75% or more of the classes compared with 79% in the original trial), meaning that many participants' dose fell short of 50 hours. The intention-to-treat finding can possibly be attributed to the patient population recruited; the general older adult, rather than selected, frailer individuals who have a history of falls, and that as a result, a more rapidly progressive FaME intervention could be used in ProAct65+. The finding from my study regarding individual dose seems inconsistent with the existing evidence. A negative correlation between exercise dose and falls was anticipated, given that dose has already been proven to be an important part of the falls prevention exercise prescription. This seemingly incongruous finding is perhaps also explained by the recruited population, of whom only 294 (24%) had fallen once or twice in the year preceding the study. 76% therefore had not fallen and many of these individuals continued to be fall-free over the duration of the therapist effect study (1063 (86%) non-fallers during the intervention, 1041 (84%) non-fallers 12 months post-intervention) irrespective of allocated study arm (327 (87%) and 388 (85%) non-fallers in FaME and UC, respectively, during the intervention and 318 (84%) and 378 (83%) non-fallers in FaME and UC, respectively, during the 12 months post-intervention). It is possible that the dose effect investigation in this study was affected by the high numbers of non-fallers.

## 6.6 Implications for further research

It is still not known whether therapist effects in falls prevention exercise exist in clinical practice, rather than in the research environment, and whether they are important predictors of patient outcome. A future observational study of 'real' therapists in practice or a trial that allows more therapist autonomy and involves less rigorous quality assurance could help elucidate this. It is interesting to note that funding has very recently been secured in the East Midlands for a study of FaME exercise in general practice to see if the effect of the intervention on falls is maintained outside of the strict research setting and to inform future commissioning of falls prevention exercise programmes (Kendrick D., personal communication, 2015). This trial will use 'fidelity' site visits to ensure PSIs are adhering to the FaME evidence base, but are not planning an exploration of therapist effects.

ProAct65+ has added to the growing body of research indicating that falls prevention interventions can be effective in the primary prevention of the first fall in the general older adult population. The population recruited were general older adults; community-dwellers, aged 65+ with a range of co-morbidities, but who were not selected according to having previously fallen. Any therapist effects may have been diluted by the number of patients who were non-fallers for the entire study duration. Some of the therapist effect sizes, although small, showed potential and it seems logical therefore to pursue the therapist effect enquiry, but within a higher risk population; possibly frequent fallers or those with a significant risk of a future fall. As well as this, falls was a secondary outcome of the ProAct65+ trial which was designed to look primarily at physical activity (PA) and longer term changes in PA behaviours, so it may also be worth conducting a therapist effect analysis in a study of FaME with falls as the primary outcome. As well as improving the detection of any therapist effects, a study of higher risk patients is clinically relevant as falls services treat those who are already falling, or who are reporting balance problems. Using a higher risk population would potentially also make the original, longer intervention desirable in future

studies for safety, and to allow sufficient time for physiological change resulting from the exercise training.

It is still not known what therapist characteristics are important in falls prevention exercise intervention. Some of the therapist characteristics investigated in this study did explain some of the small size effects. A broader range of characteristics, possibly including personality traits, could be studied in the future. As well as this, a broader investigation of what constitutes 'quality' of intervention delivered may assist in effective evaluation via 'expert' observation in clinical practice. The contribution of patient evaluation of the exercise would also benefit from further investigation.

A larger sample of therapists, each treating or allocated an equivalent number of patients, would improve the generalisability of the results and avoid allocation bias.

With regard to further investigations on dose, a purpose-designed trial would ideally have falls as the primary outcome, and rather than relying on random and differing adherence levels to 'achieve' dose variations of falls prevention exercise to analyse, patients could be allocated to groups of a pre-set dose. To allow for unplanned absences from exercise sessions (for example, illness) the length of the intervention could be flexible, thus allowing participants time to accumulate their pre-set dose.

## **6.7 Implications for training of exercise professionals**

Therapist effects in falls prevention exercise research findings could help shape future recruitment and training of exercise specialists. In the physiotherapy literature, the observation that less experienced therapists may be more capable than experienced therapists at learning and administering standardised techniques (Whitman et al., 2004) has been used to inform physiotherapy education. It has been argued that techniques which were previously reserved for higher level qualifications or continuing professional development, should be included in under-graduate curricula. Could this premise be applied to the training of exercise professionals? Should entry-level exercise qualifications include the delivery of standardised falls prevention exercises?

My study appears to support the view that experience may not be important, although the reliability of this conclusion must be viewed in light of the limitations of my exploratory study (see section 6.4). One Australian study reported improved falls and falls prevention knowledge following a falls prevention education programme delivered to university exercise science students (Pascoe et al., 2013).

## **6.8 Conclusions**

The FaME intervention appears to offer a year's 'protection' against falls beyond the end of the intervention for community-dwelling, over 65-year-olds. The shorter, more rapidly progressive FaME intervention was effective at reducing falls and fall-related injuries in this population in the ProAct65+ trial.

Therapist effects may be important in service delivery because the evidence for exercise to prevent falls has a very specific prescription, therefore maximising therapist skills and minimising 'therapist drift' could enhance patient outcomes. My exploratory study of therapist and dose effects in a large exercise trial showed no convincing evidence of either effect in relation to either falls rate or falls risk factor outcomes. This may be because quality assurance reduced variability between therapists, or because the sample size was too small to detect a small but important difference. The use of our protocol and documents for the quality assurance of the FaME intervention within research was effective at standardisation and ensuring fidelity, and this approach could be used as part of falls prevention exercise service delivery. However, the explanation that quality assurance reduces variability between therapists, as well as other possible explanations, should be tested in a purpose-designed therapist effect study.

## References

- American Geriatrics Society & British Geriatrics Society 2011. Panel on Prevention of Falls in Older Persons. Summary of the Updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. *J Am Geriatr Soc*, 59, 148–157.
- ANDERSON, T., OGLES, B. M., PATTERSON, C. L., LAMBERT, M. J. & VERMEERSCH, D. A. 2009. Therapist effects: facilitative interpersonal skills as a predictor of therapist success. *J Clin Psychol*, 65, 755-68.
- ANNESI, J. J. 1999. Relationship between exercise professionals' behavioral styles and clients' adherence to exercise. *Percept Mot Skills*, 89, 597-604.
- AREM, H., MOORE, S. C., PATEL, A., HARTGE, P., DE GONZALEZ, A. B., VISVANATHAN, K., CAMPBELL, P. T., FREEDMAN, M., WEIDERPASS, E., ADAMI, H.O. & LINET, M. S. 2015. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Internal Med*, 175, 959-967.
- BALDWIN, S. A. & IMEL, Z. E. 2013. Therapist effects: Findings and Methods. In: LAMBERT, M. J., ed. *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*. New York, Wiley, 258-297.
- BALDWIN, S. A., IMEL, Z. E. & ATKINS, D. C. 2012. The influence of therapist variance on the dependability of therapists' alliance scores: a brief comment on "The dependability of alliance assessments: the alliance-outcome correlation is larger than you think" (Crits-Christoph et al., 2011). *J Consult Clin Psychol*, 80, 947-51.
- BALDWIN, S. A., IMEL, Z. E., BRAITHWAITE, S. R. & ATKINS, D. C. 2014. Analyzing multiple outcomes in clinical research using multivariate multilevel models. *J Consult Clin Psychol*, 82, 920-30.
- BERG, K., WOOD-DAUPHINEE, S. & WILLIAMS, J. I. 1995. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand J Rehabil Med*, 27, 27-36.
- BERG, K. O., MAKI, B. E., WILLIAMS, J. I., HOLLIDAY, P. J. & WOOD-DAUPHINEE, S. L. 1992. Clinical and laboratory measures of postural balance in an elderly population. *Arch Phys Med Rehabil*, 73, 1073-80.
- BERG, K., WOOD-DAUPHINEE, S., WILLIAMS, J. I. & GAYTON, D. 1989. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Canada*, 41, 304-11.
- BLATT, S. J., SANISLOW, C. A., 3RD, ZUROFF, D. C. & PILKONIS, P. A. 1996. Characteristics of effective therapists: further analyses of data from the National

Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol*, 64, 1276-84.

BOHANNON, R. W. 2006. Reference values for the timed up and go test: a descriptive meta-analysis. *J Geriatr Phys Ther*, 29, 64-8.

BORG, G. 1982. Ratings of perceived exertion and heart rates during short-term cycle exercise and their use in a new cycling strength test. *Int J Sports Med*, 3, 153-8.

BHF (British Heart Foundation) 2015. <https://www.bhf.org.uk/publications/statistics/physical-activity-statistics-2015> (Accessed March 2016)

BROWN, A. K., LIU-AMBROSE, T., TATE, R. & LORD, S. R. 2009. The effect of group-based exercise on cognitive performance and mood in seniors residing in intermediate care and self-care retirement facilities: a randomised controlled trial. *British J Sports Med*, 43, 608-14.

BRUDER, A. M., TAYLOR, N. F., DODD, K. J. & SHIELDS, N. 2013. Physiotherapy intervention practice patterns used in rehabilitation after distal radial fracture. *Physiother*, 99, 233-40.

BUCHNER, D. M., HORN BROOK, M. C., KUTNER, N. G., TINETTI, M. E., ORY, M. G., MULROW, C. D., SCHECHTMAN, K. B., GERETY, M. B., FIATARONE, M. A. & WOLF, S. L. 1993. Development of the Common Data Base for FICSIT Trials. *J Am Geriatr Soc*, 41, 297-308.

BUTTERY, A. K., HUSK, J., LOWE, D., TREML, J., VASILAKIS, N. & RIGLIN, J. 2013. Older people's experiences of therapeutic exercise as part of a falls prevention service: survey findings from England, Wales and Northern Ireland. *Age Ageing*, 42, 182-187.

CAMPBELL, A. J., BORRIE, M. J. & SPEARS, G. F. 1989. Risk factors for falls in a community-based prospective study of people 70 years and older. *J Gerontol*, 44, M112-7.

CAMPBELL, A. J., ROBERTSON, M. C., GARDNER, M. M., NORTON, R. N. & BUCHNER, D. M. 1999a. Falls prevention over 2 years: a randomized controlled trial in women 80 years and older. *Age Ageing*, 28, 513-8.

CAMPBELL, A. J., ROBERTSON, M. C., GARDNER, M. M., NORTON, R. N. & BUCHNER, D. M. 1999b. Psychotropic medication withdrawal and a home-based exercise program to prevent falls: a randomized, controlled trial. *J Am Geriatr Soc*, 47, 850-3.

CAMPBELL, A. J., ROBERTSON, M. C., GARDNER, M. M., NORTON, R. N., TILYARD, M. W. & BUCHNER, D. M. 1997. Randomised controlled trial of a general practice programme of home based exercise to prevent falls in elderly women. *BMJ*, 315, 1065-9.

CAMPBELL, A. J., ROBERTSON, M. C., LA GROW, S. J., KERSE, N. M., SANDERSON, G. F., JACOBS, R. J., SHARP, D. M. & HALE, L. A. 2005. Randomised controlled trial of prevention of falls in people aged  $\geq 75$  with severe visual impairment: the VIP trial. *BMJ*, 331, 817.

CAMPBELL, A. J., SPEARS, G. F., BORRIE, M. J. & FITZGERALD, J. L. 1988. Falls, elderly women and the cold. *Gerontol*, 34, 205-8.

CARTER, N. D., KHAN, K. M., MCKAY, H. A., PETIT, M. A., WATERMAN, C., HEINONEN, A., JANSSEN, P. A., DONALDSON, M. G., MALLINSON, A., RIDDELL, L., KRUSE, K., PRIOR, J. C. & FLICKER, L. 2002. Community-based exercise program reduces risk factors for falls in 65- to 75-year-old women with osteoporosis: randomized controlled trial. *CMAJ*, 167, 997-1004.

Critical Appraisal Skills Programme (CASP) 1998. Qualitative research: appraisal tool. 10 questions to help you make sense of qualitative research. In Oxford: Public Health Resource Unit; 1-4. <http://www.casp-uk.net/#!/checklists/cb36> (Accessed January 2016)

Critical Appraisal Skills Programme (CASP) 2013. Systematic Review Checklist 31.05.13. <http://www.casp-uk.net/#!/checklists/cb36> (Accessed March 2016)

CELLA, M., STAHL, D., REME, S. E. & CHALDER, T. 2011. Therapist effects in routine psychotherapy practice: An account from chronic fatigue syndrome. *Psychother Res*, 21, 168-78.

CHAMBERLIN, M. E., FULWIDER, B. D., SANDERS, S. L. & MEDEIROS, J. M. 2005. Does fear of falling influence spatial and temporal gait parameters in elderly persons beyond changes associated with normal aging? *J Gerontol A Biol Sci Med Sci*, 60, 1163-7.

CLOSE, J. C. 2005a. Prevention of falls--a time to translate evidence into practice. *Age Ageing*, 34, 98-100.

CLOSE, J. C. 2005b. Prevention of falls in older people. *Disabil Rehabil*, 27, 1061-71.

CLOSE, J. C., LORD, S. L., MENZ, H. B. & SHERRINGTON, C. 2005. What is the role of falls? *Best Pract Res Clin Rheumatol*, 19, 913-35.

COTTER, P. E., TIMMONS, S., O'CONNOR, M., TWOMEY, C. & O'MAHONY, D. 2006. The financial implications of falls in older people for an acute hospital. *Irish J Med Sci*, 175, 11-3.

CRITS-CHRISTOPH, P. & MINTZ, J. 1991. Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *J Consul Clin Psychol*, 59, 20-6.

Centre for Reviews and Dissemination 2009. Systematic Reviews CRD's guidance for undertaking reviews in health care, CRD, University of York. <https://www.york.ac.uk>



/media/crd/Systematic\_Reviews.pdf (Accessed January 2016)

CUMMINGS, S. R., NEVITT, M. C. & KIDD, S. 1988. Forgetting falls. The limited accuracy of recall of falls in the elderly. *J Am Geriatr Soc*, 36, 613-6.

DEANDREA, S., LUCENTEFORTE, E., BRAVI, F., FOSCHI, R., LA VECCHIA, C. & NEGRI, E. 2010. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. *Epidemiol*, 21, 658-68.

DELBAERE, K., CLOSE, J. C., KOCHAN, N. A., SACHDEV, P. S. & LORD, S. R. 2012. Good memory as a predictor of falls: fact or artefact. *Age Ageing*, 41, 404-8.

DELBAERE, K., CLOSE, J. C., MIKOLAIZAK, A. S., SACHDEV, P. S., BRODATY, H. & LORD, S. R. 2010. The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. *Age Ageing*, 39, 210-6.

DINAN-YOUNG, S. & SKELTON, D. A., 2012. Exercise for the Prevention of Falls and Injuries in Frailer Older People (Postural Stability Instructor) Course Manual, 3rd Edition, Later Life Training.

DINAN, S. 2001. Delivering an Exercise Prescription for Vulnerable Older Patients. In YOUNG, A. & HARRIES, M. *Physical Activity for Patients: An Exercise Prescription*. London. Royal College of Physicians.

DIXON-WOODS, M., AGARWAL, S., JONES, D., YOUNG, B. & SUTTON, A. 2005. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy*, 10, 45-53.

DoH (Department of Health) 2009. Fracture Prevention Services: an economic evaluation. *dh\_110099*.

DoH (Department of Health) 2009. Falls and fractures Exercise Training to Prevent Falls. *dh\_103151*.

DoH (Department of Health) 2010. Improving care and saving money Learning the lessons on prevention and early intervention for older people. *dh\_111222*.

DoH (Department of Health) 2011. Start Active, Stay Active A report on physical activity for health from the four home countries' Chief Medical Officers. [https://www.sportengland.org/media/388152/dh\\_128210.pdf](https://www.sportengland.org/media/388152/dh_128210.pdf) (Accessed January 2016)

DOWNTON, J. H. & ANDREWS, K. 1991. Prevalence, characteristics and factors associated with falls among the elderly living at home. *Aging (Milano)*, 3, 219-28.

DUNCAN, P. W., STUDENSKI, S., CHANDLER, J. & PRESCOTT, B. 1992. Functional reach: predictive validity in a sample of elderly male veterans. *J Gerontol*, 47, M93-8.

DUNCAN, P. W., WEINER, D. K., CHANDLER, J. & STUDENSKI, S. 1990. Functional reach: a new clinical measure of balance. *J Gerontol*, 45, M192-7.

DZIEDZIC, K., HILL, J., LEWIS, M., SIM, J., DANIELS, J. & HAY, E. M. 2005. Effectiveness of manual therapy or pulsed shortwave diathermy in addition to advice and exercise for neck disorders: a pragmatic randomized controlled trial in physical therapy clinics. *Arthritis Rheum*, 53, 214-22.

EBRAHIM, S., THOMPSON, P. W., BASKARAN, V. & EVANS, K. 1997. Randomized placebo-controlled trial of brisk walking in the prevention of postmenopausal osteoporosis. *Age Ageing*, 26, 253-60.

ECCLESTONE, N. A. & JONES, C. J. 2004. International curriculum guidelines for preparing physical activity instructors or older adults. *J Ageing Phy Act*, 12, 5-21.

ENGLANDER, F., HODSON, T. J. & TERREGROSSA, R. A. 1996. Economic dimensions of slip and fall injuries. *J Forensic Sci*, 41, 733-46.

Exercise Referral Systems: A National Quality Assurance Framework 2001.  
[http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/@dh/@en/documents/digitalasset/dh\\_4079009.pdf](http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4079009.pdf) (Accessed January 2016)

FABER, M. J., BOSSCHER, R. J., CHIN, A. P. M. J. & VAN WIERINGEN, P. C. 2006. Effects of exercise programs on falls and mobility in frail and pre-frail older adults: A multicenter randomized controlled trial. *Arch Phys Med Rehabil*, 87, 885-96.

FLANNERY WAINWRIGHT S., SHEPARD, K. F. & HARMAN, J. S. 2010. Novice and experienced physical therapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. *Phys Ther*, 90, 75-88.

FLEMING, J., MATTHEWS, F. E., BRAYNE, C. and the Cambridge City over-75s Cohort study collaboration. 2008. Falls in advanced old age: recalled falls and prospective follow-up of over-90-year-olds in the Cambridge City over-75s Cohort study. *BMC Geriatr*, 8, 6.

FREEMAN, M. A. 1965. Co-ordination exercises in the treatment of functional instability of the foot. *Physiother*, 51, 393-5.

GATES, S., SMITH, L.A., FISHER, J.D. & LAMB, S.E. 2008. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. *J Rehabil Res Dev*, 45, 1105-16.

GILLESPIE, L. D., ROBERTSON, M. C., GILLESPIE, W. J., SHERRINGTON, C., GATES, S., CLEMONS, L. M. & LAMB, S. E. 2012. Interventions for preventing falls in older people living in the community. *Cochrane Database Syst Rev*, 9, CD007146.

GOLDSTEIN, H. 2011. Multilevel statistical models (Vol. 922). John Wiley & Sons.

GOODWIN, V. & BRIGGS, L. 2012. Guidelines for the physiotherapy management of older people at risk of falling. AGILE.

GRAHN KRONHED A., HALLBERG I., ODKVIST L. & MOLLER, M. 2009. Effect of training on health-related quality of life, pain and falls in osteoporotic women. *Adv Physiother*, 11, 154-65.

GREENHALGH, T. & PEACOCK, R. 2005. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*, 331, 1064-5.

GRIMLEY EVANS, J. 1997. Implications for health services. *Philos Trans R Soc Lond B Biol Sci*, 352, 1887-93.

HALTER, M., CLOSE, J., ELRICK, A., BRAIN, G. & SWIFT, C. 2000. Falls in the older population: a pilot study to assess those individuals who are attended to by the London Ambulance Service as a result of a fall but are not conveyed to an accident and emergency department. London: London Ambulance Service NHS Trust.

HANNAN, M. T., GAGNON, M. M., ANEJA, J., JONES, R. N., CUPPLES, L. A., LIPSITZ, L. A., SAMELSON, E. J., LEVEILLE, S. G. & KIEL, D. P. 2010. Optimizing the tracking of falls in studies of older participants: comparison of quarterly telephone recall with monthly falls calendars in the MOBILIZE Boston Study. *Am J Epidemiol*, 171, 1031-6.

HART, D. L. & DOBRZYKOWSKI, E. A. 2000. Influence of orthopaedic clinical specialist certification on clinical outcomes. *J Orthop Sports Phys Ther*, 30, 183-93.

HAUER, K., LAMB, S. E., JORSTAD, E. C., TODD, C. & BECKER, C. on behalf of the ProFaNE group. 2006. Systematic review of definitions and methods of measuring falls in randomised controlled fall prevention trials. *Age Ageing*, 35, 5-10.

HAWLEY-HAGUE, H., HORNE, M., CAMPBELL, M., DEMACK, S., SKELTON, D. A. & TODD, C. 2014. Multiple levels of influence on older adults' attendance and adherence to community exercise classes. *Gerontol*, 54, 599-610.

HAWLEY-HAGUE, H., HORNE, M., SKELTON, D. A. & TODD, C. 2016. Older Adults' Uptake and Adherence to Exercise Classes: Instructors' Perspectives. *J Aging Physical Activity*, 24, 119-28.

HAY, E. M., MULLIS, R., LEWIS, M., VOHORA, K., MAIN, C. J., WATSON, P., DZIEDZIC, K. S., SIM, J., MINNS LOWE, C. & CROFT, P. R. 2005. Comparison of physical treatments versus a brief pain-management programme for back pain in primary care: a randomised clinical trial in physiotherapy practice. *Lancet*, 365, 2024-30.

HEINRICH, S., RAPP, K., RISSMANN, U., BECKER, C. & KONIG, H. H. 2010. Cost of falls in old age: a systematic review. *Osteoporos Int*, 21, 891-902.

HEMMINI, E. 1982. Problems of clinical trials as evidence of therapeutic effectiveness. *Social Sci Med*, 16, 711-2.

HENDERSON, K. A. & AINSWORTH, B. E. 2000. Sociocultural perspectives on physical activity in the lives of older African American and American Indian women: a cross cultural activity participation study. *Women Health*, 31, 1-20.

HOGAN, D. B., MACDONALD, F. A., BETTS, J., BRICKER, S., EBLY, E. M., DELARUE, B., FUNG, T. S., HARBIDGE, C., HUNTER, M., MAXWELL, C. J. & METCALF, B. 2001. A randomized controlled trial of a community-based consultation service to prevent falls. *CMAJ*, 165, 537-43.

HSE (Health Survey for England) 2009. Health and lifestyles Summary of key findings. <http://www.hscic.gov.uk/catalogue/PUB00414/heal-surv-heal-life-eng-2009-rep-v1.pdf> (Accessed January 2016)

HSE (Health Survey for England) 2005. The health of older people Summary of key findings. <http://www.hscic.gov.uk/catalogue/PUB01184/heal-surv-heal-old-peo-eng-2005-rep-v5.pdf> (Accessed January 2016)

ILIFFE, S., KENDRICK, D., MORRIS, R., MASUD, T., GAGE, H., SKELTON, D., DINAN, S., BOWLING, A., GRIFFIN, M., HAWORTH, D., SWANWICK, G., CARPENTER, H.,

KUMAR, A., STEVENS, Z., GAWLER, S., BARLOW, C., COOK, J. & BELCHER, C. 2014. Multi-centre cluster randomised trial comparing a community group exercise programme with home based exercise with usual care for people aged 65 and over in primary care. *Health Technol Assess*. [http://www.ncbi.nlm.nih.gov/books/NBK262322/pdf/Bookshelf\\_NBK262322.pdf](http://www.ncbi.nlm.nih.gov/books/NBK262322/pdf/Bookshelf_NBK262322.pdf) (Accessed March 2016)

ILIFFE, S., KENDRICK, D., MORRIS, R., SKELTON, D., GAGE, H., DINAN, S., STEVENS, Z., PEARL, M. & MASUD, T. 2010. Multi-centre cluster randomised trial comparing a community group exercise programme with home based exercise with usual care for people ages 65 and over in primary care: protocol of the ProAct65+ trial. *Trials*, 11, 6.

ILIFFE, S., KENDRICK, D., MORRIS, R., GRIFFIN, M., HAWORTH, D., CARPENTER, H., MASUD, T., SKELTON, D., DINAN, S., BOWLING, A. & GAGE, H. 2015. Promoting physical activity in older people in general practice: ProAct65+ cluster randomised controlled trial. *Br J Gen Pract*, e731-8.

JADAD, A. R., MOORE, R. A., CARROLL, D., JENKINSON, C., REYNOLDS, D. J., GAVAGHAN, D. J. & MCQUAY, H. J. 1996. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*, 17, 1-12.

JELLEMA, P., VAN DER WINDT, D. A., VAN DER HORST, H. E., TWISK, J. W., STALMAN,

W. A. & BOUTER, L. M. 2005. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Cluster randomised clinical trial in general practice. *BMJ*, 331, 84.

JONES, C. J., RIKLI, R. E. & BEAM, W. C. 1999. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Res Q Exerc Sport*, 70, 113-9.

KATRAK, P., BIALOCERKOWSKI, A. E., MASSY-WESTROPP, N., KUMAR, S. & GRIMMER, K. A. 2004. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol*, 4, 22.

Kellogg International Work Group, 1987. The prevention of falls in later life. A report of the Kellogg International Work Group on the Prevention of Falls by the Elderly. *Dan Med Bull*, 34 Suppl 4, 1-24.

KEMMLER, W., VON STENGEL, S., ENGELKE, K., HABERLE, L. & KALENDER, W. A. 2010. Exercise effects on bone mineral density, falls, coronary risk factors, and health care costs in older women: the randomized controlled senior fitness and prevention (SEFIP) study. *Arch Intern Med*, 170, 179-85.

KEMPEN, G. I., YARDLEY, L., VAN HAASTREGT, J. C., ZIJLSTRA, G. A., BEYER, N., HAUER, K. & TODD, C. 2008. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. *Age Ageing*, 37, 45-50.

KENDRICK, D., KUMAR, A., CARPENTER, H., ZIJLSTRA, G. A., SKELTON, D. A., COOK, J. R., STEVENS, Z., BELCHER, C. M., HAWORTH, D., GAWLER, S. J., GAGE, H., MASUD, T., BOWLING, A., PEARL, M., MORRIS, R. W., ILIFFE, S. & DELBAERE, K. 2014. Exercise for reducing fear of falling in older people living in the community. *Cochrane Database Syst Rev*, 11, CD009848.

KERSE, N., BUTLER, M., ROBINSON, E. & TODD, M. 2004. Fall prevention in residential care: a cluster, randomized, controlled trial. *JAGS*, 52, 524-531.

KIM, D., WAMPOLD B. E. & BOLT D. M. 2006. Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychother Res*, 16(2), 161-72.

KORPELAINEN, R., KEINANEN-KIUKAANNIEMI, S., HEIKKINEN, J., VAANANEN, K. & KORPELAINEN, J. 2006. Effect of impact exercise on bone mineral density in elderly women with low BMD: a population-based randomized controlled 30-month intervention. *Osteoporos Int*, 17, 109-18.

KOSKI, K., LUUKINEN, H., LAIPPALA, P. & KIVELA, S. 1996. Physiological Factors and Medications as Predictors of Injurious Falls by Elderly People: A Prospective Population-based Study. *Age Ageing*, 25, 29-38.

KYRDALEN, I. L., MOEN, K., ROYSLAND, A. S. & HELBOSTAD, J. L. 2014. The Otago

Exercise Program performed as group training versus home training in fall-prone older people: a randomized controlled Trial. *Physiother Res Int*, 19, 108-16.

LAJOIE, Y. 2004. Effect of computerized feedback postural training on posture and attentional demands in older adults. *Aging Clin Exp Res*, 16, 363-8.

LAMB, S. E., JORSTAD-STEIN, E. C., HAUER, K. & BECKER, C. on behalf of the Prevention of Falls Network Europe and Outcomes Consensus Group. 2005. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. *J Am Geriatr Soc*, 53, 1618-22.

LEGTERS, K. 2002. Fear of falling. *Phys Ther*, 82, 264-72.

LEVSEN, M. J., HANSEN, M. L., KENT, A. D., SIEREN, J. J., THORESON, J. P. & FARRELL, K. P. 2001. Effects of Physical Therapist Training on Outcomes of Patients with Low Back Pain or Chronic Shoulder Pain. *J Man Manipul Ther*, 9, 84-91.

LEWIS, M., MORLEY, S., VAN DER WINDT, D. A., HAY, E., JELLEMA, P., DZIEDZIC, K. & MAIN, C. J. 2010. Measuring practitioner/therapist effects in randomised trials of low back pain and neck pain interventions in primary care set-tings. *Eur J Pain*, 14, 1033-9.

LI, F., FISHER, K.J., BROWNSON, R.C. & BOSWORTH, M. 2005. Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. *J Epidemiol Com Health*, 59, 558-564.

LIN, M. R., HWANG, H. F., HU, M. H., WU, H. D., WANG, Y. W. & HUANG, F. C. 2004. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. *J Am Geriatr Soc*, 52,1343–8.

LINDEMANN, U., OKSA, J., SKELTON, D. A., BEYER, N., KLENK, J., ZSCHEILE, J. & BECKER, C. 2014. Effect of cold indoor environment on physical performance of older women living in the community. *Age Ageing*, 43, 571-5.

LIU-AMBROSE, T., DONALDSON, M. G., AHAMED, Y., GRAF, P., COOK, W. L., CLOSE, J., LORD, S. R. & KHAN, K. M. 2008. Otago home-based strength and balance retraining improves executive functioning in older fallers: a randomized controlled trial. *J Am Geriatr Soc*, 56, 1821-30.

LIU-AMBROSE, T. Y., NAGAMATSU, L. S., HANDY, T. C. & LEGHARI, A. 2008. Does impaired cerebellar function contribute to risk of falls in seniors? A pilot study using functional magnetic resonance imaging. *J Am Geriatr Soc*, 56, 2153-5.

LOUGHEAD, T. M. & CARRON, A. V. 2004. The mediating role of cohesion in the leader behavior–satisfaction relationship. *Psychol Sport Exercise*, 5, 355-371.

LOGAN, P. A., COUPLAND, C. A., GLADMAN, J. R., SAHOTA, O., STONER-HOBBS, V.,

ROBERTSON, K., TOMLINSON, V., WARD, M., SACH, T. & AVERY, A. J. 2010. Community falls prevention for people who call an emergency ambulance after a fall: randomised controlled trial. *BMJ*, 340, c2102.

LORD, S. R., MARCH, L. M., CAMERON, I. D., CUMMING, R. G., SCHWARZ, J., ZOCHLING, J., CHEN, J. S., MAKAROFF, J., SITO, Y. Y., LAU, T. C., BRNABIC, A. & SAMBROOK, P. N. 2003. Differing risk factors for falls in nursing home and intermediate-care residents who can and cannot stand unaided. *J Am Geriatr Soc*, 51, 1645-50.

LORD, S. R., MCLEAN, D. & STATHERS, G. 1992. Physiological factors associated with injurious falls in older people living in the community. *Gerontol*, 38, 338-46.

LORD, S. R., WARD, J. A., WILLIAMS, P. & ANSTEY, K. J. 1993. An epidemiological study of falls in older community-dwelling women: the Randwick falls and fractures study. *Aust J Public Health*, 17, 240-5.

LORD, S. R., WARD, J. A., WILLIAMS, P. and STRUDWICK, M. 1995. The Effect of a 12 - Month Exercise Trial on Balance, Strength, and Falls in Older Women: A Randomized Controlled Trial. *J Am Geriatr Soc*, 43, 1198-206.

LORD, S. R., SHERRINGTON, C., MENZ, H. B. & CLOSE, J. C. 2007. Falls in older people: risk factors and strategies for prevention. Cambridge University Press.

LUBORSKY, L., CRITS-CHRISTOPH, P., MCLELLAN, A. T., WOODY, G., PIPER, W., LIBERMAN, B., IMBER, S. & PILKONIS, P. 1986. Do therapists vary much in their success? Findings from four outcome studies. *Am J Orthopsychiatry*, 56, 501-12.

LUCHIES, C. W., ALEXANDER, N. B., SCHULTZ, A. B. & ASHTON-MILLER, J. 1994. Stepping responses of young and old adults to postural disturbances: kinematics. *J Am Geriatr Soc*, 42, 506-12.

LUMLEY, T., DIEHR, P., EMERSON, S. & CHEN, L. 2002. The importance of the normality assumption in large public health data sets. *Annu Rev Public Health*, 23, 151-69.

MACKENZIE, L., BYLES, J. & D'ESTE, C. 2006. Validation of self-reported fall events in intervention studies. *Clin Rehabil*, 20, 331-9.

MADUREIRA, M. M., TAKAYAMA, L., GALLINARO, A. L., CAPARBO, V. F., COSTA, R. A. & PEREIRA, R. M. 2007. Balance training program is highly effective in improving functional status and reducing the risk of falls in elderly women with osteoporosis: a randomized controlled trial. *Osteoporos Int*, 18, 419-25.

MASUD, T. & MORRIS, R. O. 2001. Epidemiology of falls. *Age Ageing*, 30 Suppl 4, 3-7.  
MATHIAS, S., NAYAK, U.S. & ISAACS, B. 1986. Balance in elderly patients: the "get-up and go" test. *Arch Phys Med Rehabil*, 67, 387-9.

MEANS, K. M., RODELL, D. E. & O'SULLIVAN, P. S. 2005. Balance, mobility, and falls

among community-dwelling elderly persons: effects of a rehabilitation exercise program. *Am J Phys Med Rehabil*, 84, 238-50.

MEDELL, J. L. & ALEXANDER, N. B. 2000. A clinical measure of maximal and rapid stepping in older women. *J Gerontol A Biol Sci Med Sci*, 55, M429-33.

MIKOLAIZAK, A. S., SIMPSON, P. M., TIEDEMANN, A., LORD, S. R. & CLOSE, J. C. 2013. Systematic review of non-transportation rates and outcomes for older people who have fallen after ambulance service call-out. *Australas J Ageing*, 32, 147-57.

MOHER, D., LIBERATI, A., TETZLAFF, J. & ALTMAN, D. G. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Med*, 3, e123-30.

NANDY, S., PARSONS, S., CRYER, C., UNDERWOOD, M., RASHBROOK, E., CARTER, Y., ELDRIDGE, S., CLOSE, J., SKELTON, D., TAYLOR, S. & FEDER, G. on behalf of the Falls Prevention Pilot Steering Group. 2004. Development and preliminary examination of the predictive validity of the Falls Risk Assessment Tool (FRAT) for use in primary care. *J Public Health (Oxf)*, 26, 138-43.

NEVITT, M. C., CUMMINGS, S. R. & HUDES, E. S. 1991. Risk factors for injurious falls: a prospective study. *J Gerontol*, 46, M164-70.

NEVITT, M. C., CUMMINGS, S. R., KIDD, S. & BLACK, D. 1989. Risk factors for recurrent nonsyncopal falls. A prospective study. *JAMA*, 261, 2663-8.

NICE (National Institute for Health and Care Excellence) 2013. Falls assessment and prevention of falls in older persons. NICE Clinical Guideline 161. <http://www.nice.org.uk/CG161> (Accessed January 2016)

NICE (National Institute for Health and Care Excellence) 2012. <http://publications.nice.org.uk/the-guidelines-manual-appendices-bi-pmg6b/appendix-h-methodology-checklist-qualitative-studies> (Accessed January 2016)

NYMAN, S. R. & VICTOR, C. R. 2012. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. *Age Ageing*, 41, 16-23.

OKIISHI, J. C., LAMBERT, M. J., EGGETT, D., NIELSEN, L., DAYTON, D. D. & VERMEERSCH, D. A. 2006. An analysis of therapist treatment effects: toward providing feedback to individual therapists on their clients' psychotherapy outcome. *J Clin Psychol*, 62, 1157-72.

OKUBO, Y., SEINO, S., YABUSHITA, N., OSUKA, Y., JUNG, S., NEMOTO, M., FIGUEROA, R. & TANAKA, K. 2015. Longitudinal association between habitual walking and fall occurrences among community-dwelling older adults: analyzing the different risks of falling. *Arch Gerontol Geriatr*, 60, 45-51.



OLIVER, D., HOPPER, A. & SEED, P. 2000. Do hospital fall prevention programs work? A systematic review. *J Am Geriatr Soc*, 48, 1679-89.

OLIVO, S. A., MACEDO, L. G., GADOTTI, I. C., FUENTES, J., STANTON, T. & MAGEE, D. J. 2008. Scales to assess the quality of randomized controlled trials: a systematic review. *Phys Ther*, 88, 156-75.

O'LOUGHLIN, J. L., ROBITAILLE, Y., BOIVIN, J. F. & SUISSA, S. 1993. Incidence of and risk factors for falls and injurious falls among the community-dwelling elderly. *Am J Epidemiol*, 137, 342-54.

ONS (Office for National Statistics) 2015. <http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/compendium/opinionsandlifestylesurvey/2015-03-19/adultsmokinghabitsingreatbritain2013>. (Accessed March 2016)

PASCOE, D. A., STURNIEKS, D., CLOSE, J. C. T., TIEDEMANN, A., LORD, S., TWOMEY, D. & FINCH, C. F. 2013. Knowledge outcomes and retention of a university-based falls prevention education program (UniFPEP). *Focus Health Prof Educ*, 15, 55-66.

PATE, R. R., PRATT, M., BLAIR, S. N., HASKELL, W. L., MACERA, C. A., BOUCHARD, C., BUCHNER, D., ETTINGER, W., HEATH, G. W. & KING, A. C. 1995. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA*, 273, 402-7.

PERRY, L., KENDRICK, D., MORRIS, R., DINAN, S., MASUD, T., SKELTON, D. & ILIFFE, S. for the ProAct65+ Study Team. 2012. Completion and return of fall diaries varies with participants' level of education, first language, and baseline fall risk. *J Gerontol A Biol Sci Med Sci*, 67, 210-4.

PODSIADLO, D. & RICHARDSON, S. 1991. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*, 39, 142-8.

POPAY, J., ROBERTS, H., SOWDEN, A., PETTICREW, M., ARAI, L., RODGERS, M., BRITTEN, N., ROEN, K., DUFFY, S. 2006. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews A Product from the ESRC Methods Programme.

POPE, C., MAYS, N. & POPAY, J. 2006. How can we synthesize qualitative and quantitative evidence for healthcare policy-makers and managers? *Healthc Manage Forum*, 19, 27-31.

POWELL, L. E. & MYERS, A. M. 1995. The activities-specific balance confidence (ABC) scale. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 50, 28-34.

RAND, D., MILLER, W. C., YIU, J., & ENG, J. J. 2011. Interventions for addressing low

balance confidence in older adults: a systematic review and meta-analysis. *Age Ageing*, 40, 297-306.

RASMUSSEN, K. & ANDERSEN, J. H. 2005. Individual factors and GP approach as predictors for the outcome of rehabilitation among long-term sick listed cases. *J Occup Rehabil*, 15, 227-35.

RESNIK, L. & HART, D. L. 2003. Using clinical outcomes to identify expert physical therapists. *Phys Ther*, 83, 990-1002.

RESNIK, L. & HART, D. L. 2004. Influence of Advanced Orthopaedic Certification on Clinical Outcomes of patients with Low Back Pain. *J Man Manipul Ther*, 12, 32-43.

RESNIK, L. & JENSEN, G. M. 2003. Using clinical outcomes to explore the theory of expert practice in physical therapy. *Phys Ther*, 83, 1090-106.

ROBERTSON, M. C., CAMPBELL, A. J. & HERBISON, P. 2005. Statistical analysis of efficacy in falls prevention trials. *J Gerontol A Biol Sci Med Sci*, 60, 530-4.

ROBERTSON, M. C., DEVLIN, N., GARDNER, M. M. & CAMPBELL, A. J. 2001a. Effectiveness and economic evaluation of a nurse delivered home exercise programme to prevent falls. 1: Randomised controlled trial. *BMJ*, 322, 697-701.

ROBERTSON, M. C., GARDNER, M. M., DEVLIN, N., MCGEE, R. & CAMPBELL, A. J. 2001b. Effectiveness and economic evaluation of a nurse delivered home exercise programme to prevent falls. 2: Controlled trial in multiple centres. *BMJ*, 322, 701-4.

ROGERS, M. W., HEDMAN, L. D., JOHNSON, M. E., MARTINEZ, K. M. & MILLE, M. L. 2003. Triggering of protective stepping for the control of human balance: age and contextual dependence. *Brain Res Cogn Brain Res*, 16, 192-8.

ROSE, D. J., JONES, C. J. & LUCCHESI, N. 2002. Predicting the probability of falls in community-residing older adults using the 8-foot up-and-go: a new measure of functional mobility. *J Aging Phys Act*, 10, 466-75.

ROSSITER-FORNOFF, J. E., WOLF, S. L., WOLFSON, L. I. & BUCHNER, D. M. 1995. A cross-sectional validation study of the FICSIT common data base static balance measures. *Frailty and Injuries: Cooperative Studies of Intervention Techniques*. *J Gerontol A Biol Sci Med Sci*, 50, M291-7.

RCP (Royal College of Physicians) 2012. Older people's experiences of therapeutic exercise as part of a falls prevention service - patient and public involvement. London: RCP.

RUBENSTEIN, L. Z. 2006. Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing*, 35 Suppl 2, ii37-ii41.

SAUCIER, G. 1994. Separating description and evaluation in the structure of personality attributes. *J Pers Soc Psychol*, 66, 141-54.

SEGUIN, R. A., ECONOMOS, C. D., PALOMBO, R., HYATT, R., KUDER, J. & NELSON, M. E. 2010. Strength training and older women: a cross-sectional study examining factors related to exercise adherence. *J Aging Phys Act*, 18, 201-18.

SCHEFFER, A. C., SCHUURMANS M. J., VAN DIJK, N., VAN DER HOOFT, T. & DE ROOIJ, S. E. 2008. Fear of falling: measurement strategy, prevalence, risk factors and consequences among older persons. *Age Ageing*, 37, 19-24.

SCHOENE, D., WU, S. M., MIKOLAIZAK, A. S., MENANT, J. C., SMITH, S. T., DELBAERE, K. & LORD, S. R. 2013. Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta-analysis. *J Am Geriatr Soc*, 61, 202-8.

SCOTT, V., VOTOVA, K., SCANLAN, A. & CLOSE, J. 2007. Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. *Age Ageing*, 36, 130-9.

SCUFFHAM, P., CHAPLIN, S. & LEGOOD, R. 2003. Incidence and costs of unintentional falls in older people in the United Kingdom. *J Epidemiol Community Health*, 57, 740-4.

SHERRINGTON, C., TIEDEMANN, A., FAIRHALL, N., CLOSE, J. C. & LORD, S. R. 2011. Exercise to prevent falls in older adults: an updated meta-analysis and best practice recommendations. *N S W Public Health Bull*, 22, 78-83.

SHERRINGTON, C., WHITNEY, J. C., LORD, S. R., HERBERT, R. D., CUMMING, R. G. & CLOSE, J. C. 2008. Effective exercise for the prevention of falls: a systematic review and meta-analysis. *J Am Geriatr Soc*, 56, 2234-43.

SHUMWAY-COOK, A., BRAUER, S. & WOOLLACOTT, M. 2000. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther*, 80, 896-903.

SIMON, C. B., STRYKER, S. E. & GEORGE, S. Z. 2013. Assessing the influence of treating therapist and patient prognostic factors on recovery from axial pain. *J Man Manip There*, 21, 187-95.

SIMPSON, J. M., WORSFOLD, C., FISHER, K. D. & VALENTINE, J. D. 2009. The CONFbal scale: a measure of balance confidence--a key outcome of rehabilitation. *Physiother*, 95, 103-9.

SIMPSON, J. M., WORSFOLD, C. & HAWKE, J. 1998. Balance Confidence in Elderly People the CONFbal Scale. *Age Ageing*, 27, 57.

SKELTON, D. A., ASTON, H., GREIG, C. A. & YOUNG, A. 1992. Effect of Cooling on the

Power of Voluntary and Electrically Stimulated Contractions of the Lower Limb Extensors. *Clinical Sci* 83, 16.

SKELTON, D., DINAN, S., CAMPBELL, M. & RUTHERFORD, O. 2005. Tailored group exercise (Falls Management Exercise - FaME) reduces falls in community-dwelling older frequent fallers (an RCT). *Age Ageing*, 34, 636-9.

SKELTON, D. A. & DINAN, S. 1999. Exercise for falls management: Rationale for an exercise programme aimed at reducing postural instability. *Physiother Theory Pract*, 15, 105-20.

SONG, J. W. & CHUNG, K. C. 2010. Observational Studies: Cohort and Case-Control Studies. *Plast Reconstr Surg*, 126, 2234-42.

SPENCER, L., RITCHIE, J., LEWIS, J. & DILLON, L. 2003. Quality in qualitative evaluation: a framework for assessing research evidence [monograph online]. London: Cabinet Office. <http://dera.ioe.ac.uk/21069/2/a-quality-framework-tcm6-38740.pdf> (accessed January 2016)

STEVENS, J. A., BALLESTEROS, M. F., MACK, K. A., RUDD, R. A., DECARO, E. & ADLER, G. 2012. Gender differences in seeking care for falls in the aged Medicare population. *Am J Prev Med*, 43, 59-62.

STEVENS, J. A., THOMAS, K. E. & SOGOLOW, E. D. 2007. Seasonal patterns of fatal and nonfatal falls among older adults in the U.S. *Accid Anal Prev*, 39, 1239-44.

STEVENS, Z., CARPENTER, H., GAWLER, S., BELCHER, C., HAWORTH, D., KENDRICK, D., MORRIS, R., MASUD, T., SKELTON, D. A. & ILIFFE, S. 2013. Lessons learnt during a complex, multicentre cluster randomised controlled trial: the ProAct65+ trial. *Trials*, 14, 192.

THOMAS, J. & HARDEN, A. 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Research Methodol*, 8, 1-19.

THOMAS, S., MACKINTOSH, S. & HALBERT, J. 2010. Does the 'Otago exercise programme' reduce mortality and falls in older adults?: a systematic review and meta-analysis. *Age Ageing*, 39, 681-7.

TINETTI, M. E. 1986. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc*, 34, 119-26.

TINETTI, M. E. 2010. Making prevention recommendations relevant for an aging population. *Ann Intern Med*, 153, 843-4.

TINETTI, M. E. & KUMAR, C. 2010. The patient who falls: "It's always a trade-off". *JAMA*, 303, 258-66.

- TINETTI, M. E., LIU, W. L. & CLAUS, E. B. 1993. Predictors and prognosis of inability to get up after falls among elderly persons. *JAMA*, 269, 65-70.
- TINETTI, M. E., SPEECHLEY, M. & GINTER, S. F. 1988. Risk factors for falls among elderly persons living in the community. *N Engl J Med*, 319, 1701-7.
- TINETTI, M. E., WILLIAMS, T. F. & MAYEWSKI, R. 1986. Fall risk index for elderly patients based on number of chronic disabilities. *Am J Med*, 80, 429-34.
- TODD, C. & SKELTON, D. 2004. What are the Main Risk Factors for Falls Amongst Older People and what are the Most Effective Interventions to Prevent These Falls? World Health Organisation.
- TORGERSON, D. J. & BELL-SYER, S. E. 2001. Hormone replacement therapy and prevention of nonvertebral fractures: a meta-analysis of randomized trials. *JAMA*, 285, 2891-7.
- VAN DER MARCK, M. A., OVEREEM, S., KLOK, P. C., BLOEM, B. R. & MUNNEKE, M. 2011. Evaluation of the falls telephone: an automated system for enduring assessment of falls. *J Am Geriatr Soc*, 59, 340-4.
- VAN DIJK, P. T., MEULENBERG, O. G., VAN DE SANDE, H. J. & HABBEMA, J. D. 1993. Falls in dementia patients. *Gerontologist*, 33, 200-4.
- VOUKELATOS, A., CUMMING, R. G., LORD, S. R. & RISSEL, C. 2007. A randomized, controlled trial of tai chi for the prevention of falls: the Central Sydney tai chi trial. *J Am Geriatr Soc*, 55, 1185-91.
- WALLER, G. 2009. Evidence-based treatment and therapist drift. *Behav Res Ther*, 47, 119-27.
- WALWYN, R. & ROBERTS, C. 2010. Therapist variation within randomised trials of psychotherapy: implications for precision, internal and external validity. *Stat Methods Med Res*, 19, 291-315.
- WAMPOLD, B. E., IMEL, Z. E., LASKA, K. M., BENISH, S., MILLER, S. D., FLUCKIGER, C., DEL RE, A. C., BAARDSETH, T. P. & BUDGE, S. 2010. Determining what works in the treatment of PTSD. *Clin Psychol Rev*, 30, 923-33.
- WAMPOLD, B. E. & SERLIN, R. C. 2000. The consequence of ignoring a nested factor on measures of effect size in analysis of variance. *Psychol Methods*, 5, 425-33.
- WESTFALL, P. H. & YOUNG, S. S. 1993. On Adjusting P-Values for Multiplicity. *Biometrics*, 49, 941-4.
- WHITE, S.M. & GRIFFITHS, R. 2011. Projected incidence of proximal femoral fracture in England: a report from the NHS Hip Fracture Anaesthesia Network (HIPFAN). *Injury*

42, 1230-1233.

WHITMAN, J. M., FRITZ, J. M. & CHILDS, J. D. 2004. The influence of experience and specialty certifications on clinical outcomes for patients with low back pain treated within a standardized physical therapy management program. *J Orthop Sports Phys Ther*, 34, 662-72; discussion 672-5.

WHITNEY, J. C., LORD, S. R. & CLOSE, J. C. 2005. Streamlining assessment and intervention in a falls clinic using the Timed Up and Go Test and Physiological Profile Assessments. *Age Ageing*, 34, 567-71.

WILD, D., NAYAK, U. S. & ISAACS, B. 1981. How dangerous are falls in old people at home? *Br Med J (Clin Res Ed)*, 282, 266-8.

WOLF, S. L., BARNHART, H. X., KUTNER, N. G., MCNEELY, E., COOGLER, C. & XU, T. 1996. Reducing frailty and falls in older persons: an investigation of Tai Chi and computerized balance training. Atlanta FICSIT Group. Frailty and Injuries: Co-operative Studies of Intervention Techniques. *J Am Geriatr Soc*, 44, 489-97.

World Health Organisation 2007. WHO Global Report on Falls Prevention in Older Age. Geneva (Switzerland). [http://www.who.int/ageing/publications/Falls\\_prevention7March.pdf](http://www.who.int/ageing/publications/Falls_prevention7March.pdf) (Accessed March 2016)

YARDLEY, L., BEYER, N., HAUER, K., KEMPEN, G., PIOT-ZIEGLER, C. & TODD, C. 2005. Development and initial validation of the Falls Efficacy Scale-International (FES-I). *Age Ageing*, 34, 614-9.

ZECEVIC, A. A., SALMONI, A. W., SPEECHLEY, M. & VANDERVOORT, A. A. 2006. Defining a fall and reasons for falling: comparisons among the views of seniors, health care providers, and the research literature. *Gerontologist*, 46, 367-76.

### STUDY PROTOCOL

### Open Access

# Multi-centre cluster randomised trial comparing a community group exercise programme with home based exercise with usual care for people aged 65 and over in primary care: protocol of the ProAct 65+ trial

Steve Iliffe<sup>1\*</sup>, Denise Kendrick<sup>2</sup>, Richard Morris<sup>1</sup>, Dawn Skelton<sup>3</sup>, Heather Gage<sup>4</sup>, Susie Dinan<sup>1</sup>, Zoe Stevens<sup>1</sup>, Mirilee Pearl<sup>1</sup>, Tahir Masud<sup>5</sup>

#### Abstract

**Background:** Regular physical activity reduces the risk of mortality from all causes, with a powerful beneficial effect on risk of falls and hip fractures. However, physical activity levels are low in the older population and previous studies have demonstrated only modest, short-term improvements in activity levels with intervention.

**Design/Methods:** Pragmatic 3 arm parallel design cluster controlled trial of class-based exercise (FAME), home-based exercise (OEP) and usual care amongst older people (aged 65 years and over) in primary care. The primary outcome is the achievement of recommended physical activity targets 12 months after cessation of intervention. Secondary outcomes include functional assessments, predictors of exercise adherence, the incidence of falls, fear of falling, quality of life and continuation of physical activity after intervention, over a two-year follow up. An economic evaluation including participant and NHS costs will be embedded in the clinical trial.

**Discussion:** The ProAct65 trial will explore and evaluate the potential for increasing physical activity among older people recruited through general practice. The trial will be conducted in a relatively unselected population, and will address problems of selective recruitment, potentially low retention rates, variable quality of interventions and falls risk.

**Trial Registration:** Trial Registration: ISRCTN43453770

#### Background

The health benefits of physical activity have been extensively reviewed and evidence suggests that it reduces the risk of cardiovascular disease, type 2 diabetes, osteoporosis and certain cancers [1]. There is growing evidence of the association between regular physical activity and a reduced risk of all cause mortality [2], and of the potential savings for NHS budgets from exercise promotion for older adults [3]. Sedentary behaviour increases the risk of dependence, falls and fractures. Sustained levels of physical activity in adulthood

maintain bone strength and can prevent fragility fractures in later life. Research has shown that a lifetime's history of regular physical activity can reduce the risk of hip fracture by up to 50% and much of this benefit is thought to result from a reduction in falls [4]. It is now clear that habitual physical activity and improved access to exercise opportunities is an important public health approach to the prevention of functional decline that can lead to frailty, falls and fractures [5].

Falls are common in people aged 65 years and older and can have serious consequences, including injury, pain, impaired function, loss of confidence in carrying out everyday activities, loss of independence and autonomy, and even death [6,7]. There is evidence that

\* Correspondence: [s.iliffe@pcps.ucl.ac.uk](mailto:s.iliffe@pcps.ucl.ac.uk)

<sup>1</sup>Department of Primary Care & Population Health, University College London, Rowland Hill St, London NW3 2PF, UK

interventions providing some form of exercise may be effective in preventing falls amongst older people [8] and that healthcare costs [9,10] can be reduced if falls are reduced [11-15].

Current recommendations for health benefits are that people do at least 30 minutes of physical activity of moderate intensity on at least five days of the week [16]. However, surveys have consistently shown a high prevalence of physical inactivity in the UK population [17]. A recent systematic review comparing seventeen randomised controlled trials with different interventions designed to encourage sedentary, community dwelling adults to do more physical activity [18] concluded that interventions were effective in the short and mid term, at least in middle age, and that there were no significant increases in adverse events in the four studies that reported them. However, it is unclear which individual interventions (e.g. home-based or facility-based) are the most effective in increasing physical activity in the long term, or in specific groups (e.g. older people).

The NHS is attempting to address the problem of inactivity in a variety of ways, including exercise referral schemes in primary care ('exercise on prescription') which are currently provided by approximately 89% of Primary Care Trusts and usually involves referring patients to local leisure centres [19]. Although exercise on prescription has been shown to be feasible and effective in vulnerable older people [20], there appear to be significant barriers to the uptake of exercise classes in leisure centres. For many older people home exercise or group exercise in non-intimidating environments (e.g. community halls) may be more appealing, and result in higher uptake of exercise programmes and longer continuation of exercise. Peer activity mentors have also been shown to be effective in increasing uptake and adherence to exercise [21-24].

There are currently two existing exercise programmes, designed for use in community settings, specifically for people aged 65 and over. The first is a home based programme known as the Otago Exercise Programme (OEP) and the second is a community based group exercise programme known as the Falls Management Exercise Programme (FaME).

The OEP (Otago Exercise Programme) [25-31] and FaME (Falls Management Exercise) programmes [32] were both designed for use in community settings, specifically for people aged 65 and over. As well as being designed to reduce falls, both are based on the components of fitness and principles of programming for all older adults (i.e. warm up, mobility, stretches, strength and balance, endurance and a cool-down) and have all the elements of training appropriate for that age group. Exercises are tailored to the individual's ability and health need. Both programmes involve strength &

balance training which is tailored to the individual's ability and health need.

The OEP is a home based exercise programme for older people which is effective in reducing the falls and fall-related injuries, improving balance, strength and confidence in performing everyday activities without falling, and has been shown to be cost effective for people aged 80 and over [25-31]. It was designed to be delivered by physiotherapists and nurses trained and supervised by physiotherapists. A one year evaluation of the OEP showed considerable improvements in outdoor activities (walking, shopping, gardening, other outside leisure activities) after 6 months (unpublished data Campbell) with participants continuing to exercise after completing the programme. It also showed significant improvements in executive function after 6 months [30]. Whilst the OEP has been evaluated in four controlled trials of older primary care patients in New Zealand and one RCT in Canada, it has not been tested in a primary care setting in the UK for its feasibility, impact, acceptability and cost-effectiveness.

FaME is a group exercise programme which was developed and tested in a controlled trial in the UK [32], but not in a primary care population. It aims to improve balance [33] and was designed to be delivered by qualified postural stability instructors (PSIs) [34]. It has been shown to be effective in reducing falls and injuries resulting from falls [32]. Good compliance was demonstrated with the FaME programme and nearly two thirds of people participating in FaME continued in group exercise programmes for over a year after trial completion. (unpublished data Skelton). FaME now needs to be evaluated for its impact, acceptability and cost-effectiveness within primary care.

This trial aims to fill the gaps in the current evidence base by evaluating the delivery, impact, acceptability and cost-effectiveness of a community based exercise programme (FaME) and a home based exercise programme (OEP) supported by similarly aged (peer) mentors, compared with usual care for primary care patients. The underlying assumption is that the exercises will produce sufficient subjective well-being and improved mobility to encourage continuation of higher levels of physical activity after the cessation of the intervention. Each exercise programme will be compared with usual care for effectiveness in producing sustained change in physical activity. The two programmes will be compared for cost-effectiveness if both are effective in promoting sustained change in physical activity. The primary hypotheses are: 1) Both exercise programmes produce sustained changes in physical activity (PA) compared with usual care, and 2) The Otago programme (OEP) is more cost effective than FaME.



## Objectives

The primary objective is to determine the effect of two evidence based exercise programmes designed for older people, compared with usual care (i.e. with no special interventions to promote physical activity) on the achievement of recommended physical activity targets 12 months after cessation of intervention.

The secondary objectives are to:

(1) determine the health benefits of the programmes to participants starting at various levels of physical activity, particularly the effects on physical and psychological status, health status, health-related quality of life and quality adjusted life years (QALYs).

(2) estimate the costs of the exercise interventions, and possible cost offsets, and to assess the cost-effectiveness of community group exercise, and home-supported exercise compared with each other, and with usual care.

(3) determine the acceptability of the programmes, adherence rates, enabling factors and barriers to future implementation.

(4) compare the time course of responses by participants in terms of exercising at the recommended levels, at 0, 6, 12, 18 and 24 months after cessation of the intervention, between those undergoing the exercise programmes, and those receiving usual care.

(5) determine participants' perceptions of the value of exercise, and the predictors of continued exercise.

## Design/Methods

A 3 arm parallel design cluster controlled trial using minimisation for allocation at the level of general practice in two centres (London and Nottingham/Derby), comparing community-centre based group exercise programme (FaME), with a home based exercise programme and walking plan (OEP) and with usual care. There will be two years' follow-up to determine the impact, acceptability and adherence to the programme, longer term continuation of exercise and cost-effectiveness. The CONSORT diagram [35] summarises the design (Fig 1).

### Participants

Participants will be patients aged 65 years and over registered with participating general practices who give informed consent to participate.

### Inclusion criteria for practices

Inclusion criteria will be a commitment to participate over the duration of the study and the availability of a suitable community venue in the practice catchment area.

### Inclusion criteria for participants

Those aged 65 years and older, who can walk around independently of personal help both indoors and outdoors (with or without a walking aid) and who would be physically able to take part in a group exercise class, will be eligible to participate.

### Exclusion criteria for participants

Those with any of the following criteria will be excluded:

- Three or more self-reported falls in the previous year;
- Resting BP > 180/100 mmHg; tachycardia > 100 bpm; those considered by their GP to have uncontrolled hypertension; significant drop in BP during exercise recorded in the medical records or found at initial assessment;
- Psychiatric conditions which would prevent participation in an exercise class, e.g. psychotic illness;
- Uncontrolled medical problems, which the GP considers would exclude patients from undertaking the exercise programme; e.g. acute systemic illness' such as pneumonia, poorly controlled angina, acute rheumatoid arthritis, unstable or acute heart failure;
- Conditions requiring a specialist exercise programme, e.g. uncontrolled epilepsy, significant neurological disease or impairment; unable to maintain seated upright position or unable to move about independently indoors;
- Not living independently (e.g. residential or nursing care);
- Significant cognitive impairment (unable to follow simple instructions);
- Already receiving long term physiotherapy or already in an exercise programme.

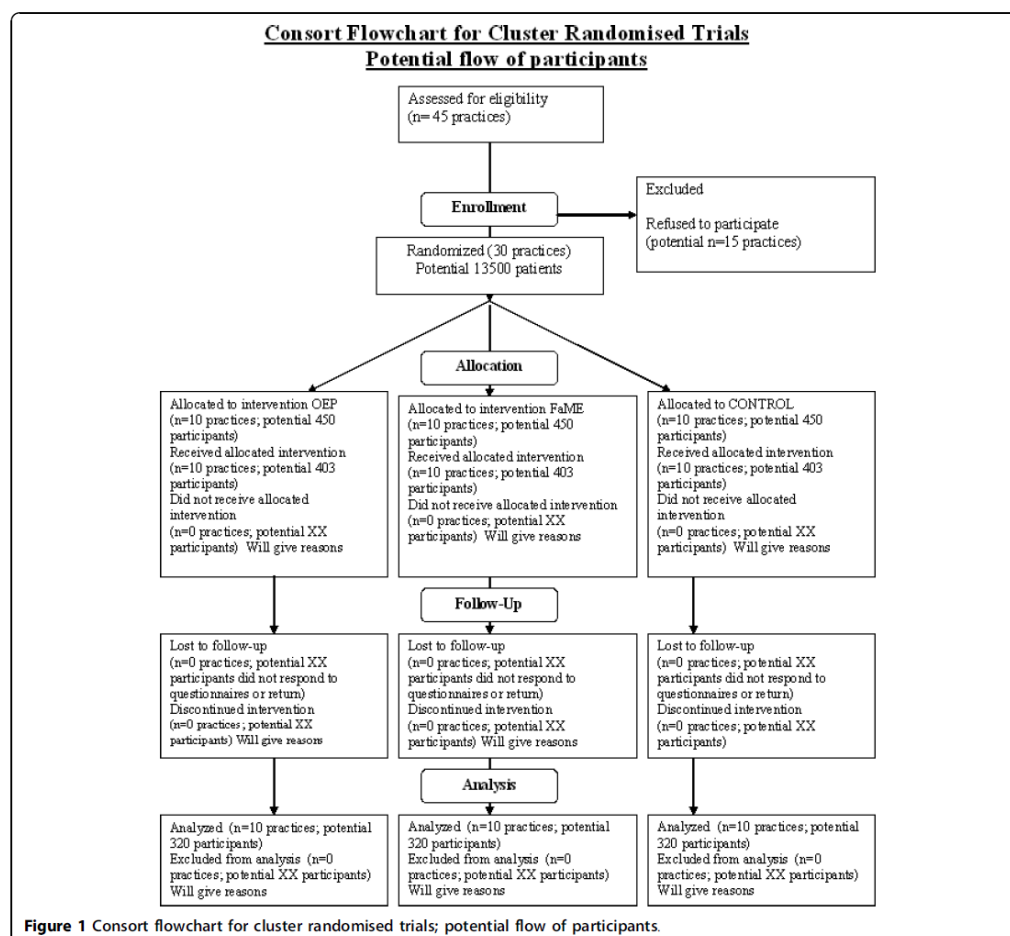
Exclusion criteria will be checked at the recruitment appointment by the researcher. This will include measurement of resting BP and pulse and completion of a health questionnaire. The participant's GP will be asked to confirm eligibility for each potentially eligible participant.

### Recruitment of practices

General practices will be recruited from the Primary Care Research Networks (PCRN) in London and Nottingham/Derby. The PCRN's will be asked to identify potential participant practices by size and deprivation status. Approaches by mailed invitation, telephone contact with practice managers, and personal contact with local GP opinion leaders all be used as necessary [36].

### Recruitment of participants

Practices will produce a single numbered list of patients aged 65 and over. Practice staff will be allowed to make and justify their own exclusions in liaison with the research team. The research team will provide the practices with a random number list to select the sample of patients to be approached after exclusions have been made. The sampling will vary depending on practice size. In practices with fewer than 600 patients aged 65 and over, all patients aged 65 and over will be invited to participate. In larger practices random sampling will be used to identify 600 patients aged 65 and over who will



be invited to participate. Patients will then be sent invitation letters about the trial by their usual General Practitioner.

#### Interventions

There are 3 arms to the trial:

- (1) home based exercise programme and walking plan (OEP)
- (2) community-centre based group exercise programme (FaME)
- (3) usual care

#### Home based exercise programme (OEP)

This consists of a 30 minute programme of leg muscle strengthening and balance retraining exercises progressing in difficulty to be performed at home at least three times per week, and a walking plan to be undertaken at

least two times per week for 24 weeks. Each participant will receive an instruction booklet and ankle cuff weights (starting at 0.5 kg) to provide resistance for the strengthening exercises. The programme will be introduced to participants by trained research staff, at an appropriate starting level determined at an initial assessment. Trained peer mentors will contact and visit the participants at home to start the exercise programme with them and will follow-up with up to three more home visits (as the participants require). Participants will record the days they complete the programme, and mentors will telephone them fortnightly as mentor support has been shown to be effective in increasing adherence [21-23]. Mentors will record and report any problems encountered with the exercise programme to

the research team using an adverse event form developed for the study.

The delivery of the OEP will be standardised through training of research staff before the trial starts and there will be regular contact with the participants and peer mentors to check delivery protocols are being followed.

#### **Community based exercise programme (FaME)**

The FaME programme comprises one hour-long PSI-delivered group exercise class in a local community centre for a maximum of 15 participants, and two 30 minute home exercise sessions (based on the OEP) per week for 24 weeks. Participants will be advised to walk at least twice per week for up to 30 minutes at a moderate pace. The programme includes leg muscle strengthening and balance retraining that progress in difficulty. Progressive trunk and arm muscle strengthening, bone loading, endurance (including walking) and flexibility training, functional floor skills (see below) and adapted Tai Chi complete the evidence based programme. Ankle cuff weights, Therabands (elastic resistance training bands) and mats are also used throughout the programme. The group exercises include retraining of the ability to get up from the floor and floor exercises to improve strength, balance and coping strategies to reduce the risk of complications resulting from a long lie [34].

The delivery of the FaME programme will be standardised through training of PSIs before the trial starts and there will be regular quality assurance visits for the FaME classes to check delivery protocols are being followed.

The PSI will keep a register of attendance and record tailoring of the programme and any feedback from participants. They will follow up non-attenders by telephone as necessary, recording any positive or negative feedback, and notify the research team about reasons for non-attendance or drop-out. Participants will be given a booklet containing their home exercise instructions.

FaME groups will contain 9 or 10 participants, so there will be 4 or 5 classes per week for each of the practices allocated to this arm. The number of PSIs running these classes will be determined by their availability, but the aim is to maximise continuity and standardisation in physical activity training, so the ideal would be to have one PSI who would lead all groups in one practice. This may be difficult to achieve, but we will report the actual deployment of PSIs in our findings. The same applies to peer mentors for OEP classes.

General practitioners in participating practices allocated to either the FaME or OEP programmes will be discouraged from referring participants involved in the trial to other exercise therapy projects, outside the study.

#### **Usual care**

Participants in the usual care arm will not be offered either the OEP or FaME programmes. They will be free

to participate in any other exercise as they would if they were not participating in the trial.

#### **Cultural and ethnic sensitivity**

Cultural and religious requirements will be accommodated within the exercise programmes. The recommendations from the Help the Aged Minority Ethnic Elders Falls Prevention Programme (MEEFP: <http://www.help-theaged.org.uk/meefp>) will be followed and the research team will liaise closely with Skills Active (Sector Skills Council for Active Leisure and Learning), who are working with the Integrated Fitness Initiative's (IFI) current programme: Physical Activity Provision for Ethnic Minority Groups. In particular the FaME group classes will ensure that recommendations for attire will respect cultural, religious beliefs and customs for a range of ethnic groups. Single sex exercise groups will be scheduled as required, and separate changing facilities and gender specific instructors will be provided wherever possible. Windows in the exercise classrooms will be screened as appropriate. Family support will be encouraged, and classes will be provided at different times of the day. The OEP programme will respect participant's preferences regarding family support and participation in the home exercise programme.

All research material and exercise manuals will use a maximum reading age of 9 years. Inability to read the material is not a formal exclusion criterion as the individual may be able to follow movement and correction accurately in classes and family members will be allowed to act as interpreters. Where possible, translations of research material and exercise manuals will be provided.

#### **Outcome measures**

The primary and secondary outcome measures have been chosen to reflect the needs of participants (e.g. functional outcomes, falls, confidence, quality of life, participant costs), of commissioners of exercise services in primary care and policy makers (e.g. physical activity, falls, NHS costs).

The primary outcome will be the proportion reaching the recommended physical activity (PA) target of at least 30 minutes of activity of moderate intensity on at least 5 days each week, measured using the CHAMPS, PASE and PHONE\_FITT questionnaires. While measures will be taken at 0, 6, 12, 18 and 24 months after the intervention, our primary analysis will be of 12 month data, as this is the time when the effect of the intervention is expected to be maximal, as seen in a previous evaluation of the OEP in New Zealand [37].

The secondary outcomes will include:

(1) The direct health benefits: i.e. functional and psychological status, the rate of falls (the major safety outcome measure), the number and nature of falls, and fear of falling,



(2) Stage of change, self-efficacy for exercise and physical self-perception (self-esteem relative to the physical domain), which includes measurement of perceived importance (the degree to which participants value their physical condition, body image and physical strength) to inform predictors of exercise adherence and continuation, and participants' judgement of the value or importance of physical activity,

(3) Health - related quality of life and Quality adjusted life years (QALYs) [38],

(4) The NHS and private (participant) costs of each exercise programme, and possible cost offsets, identified from a comparison of health and social service utilisation of participants in all groups during the study period.

#### **Ascertainment of outcomes**

The following functional assessments will be used by researchers at baseline and at the end of the interventions (and at 6 months after allocation in the usual care arm):

1. Modified Clinical Romberg Static Balance test, eyes open and closed [39].
2. Timed get-up and go (with and without distraction) (TUG) as a measure of balance and falls risk [40].
3. Functional Reach as a measure of balance and falls risk [41].
4. 30 second chair rise as a measure of lower limb strength and power [42].

The following validated tools will be used at baseline, and as self-completion questionnaires at follow-up:

1. Confidence in balance measured by the ConfBal scale [43]. A total score is provided as a measure of confidence.
2. Confidence in carrying out a range of basic activities of daily living without falling measured by the Falls Efficacy Scale-International (FES-I) [44].
3. Readiness to change measured by the transtheoretical model [45] and applying it to exercise behaviour to determine perceived barriers [46] and self efficacy for exercise [47].
4. Quality of life will be measured using Older People's QoL Questionnaire (OPQOL) [48-50]
5. Social network size and density will be measured using the brief Lubben Social Network scale [51] and perceived social support measured by the Multidimensional Scale of Perceived Social Support (MSPSS) [52].
6. Subjective Habitual Physical Activity will be assessed using a number of validated questionnaires to ensure all domains of activity and sport are considered, including the Phone-FITT, PASE and CHAMPS [53-55] and the current level of activity questions used in the Household Survey [56].
7. Attitudes and beliefs about falls prevention interventions will be measured using the AFRIS questionnaire [57].

8. Falls risk will be measured by the Falls Risk Assessment Tool (FRAT) [58].

9. Health- related quality of life will be measured by the SF-12 [59]. Quality Adjusted Life Years (QALYs), which are the main outcome for the economic analysis, will be based on EQ-5D utility weights obtained by transforming SF-12 scores [38].

In addition, demographic information, co-morbidity, medication, use of general practice and hospital and community social services will also be recorded at baseline and updated at subsequent assessments. Falls will be ascertained by self-completed fall diaries, with follow up of non-responders and telephone contact with fallers to ascertain the type of fall and any injury and health care usage that resulted.

Feedback will be sought from all exercise participants using a questionnaire which will include open questions eliciting views of the programme, reasons for drop-out, barriers to attendance and self perceptions of the benefits and disadvantages of the programmes to aid future implementation.

For the purposes of the economic analysis, the resources used in the delivery of the interventions will be collected from records kept by PSI instructors (FaME) and the research staff and peer mentors (OEP). The use of facilities and equipment, and the time spent on travel and instruction will be included and monetary costs will be assigned according to market rates.

In addition, the use of health and social care services (GP, community, outpatient, hospital admission) will be recorded for participants in all groups by means of the falls diaries. Self reported service utilisation will be verified from the primary care medical records of consenting patients after the follow up period. Costs of services will be obtained from local and national sources [60]. Health and social care costs in the exercise groups will be compared with each other and with the usual care (no exercise) group to assess the extent to which the costs of the exercise intervention may be offset by savings elsewhere in the health and social care system.

No other encouragement to continue with physical activity will be given to participants, and all potential reinforcements in the form of diaries and six monthly contracts will be given to all three arms of the trial.

#### **Baseline data collection**

Baseline assessment will include all functional assessments plus administration of all questionnaires described above.

#### **Follow-up data collection**

Follow up assessments occur at 24 weeks after the commencement of the intervention, and at 6, 12, 18 and 24 months after the completion of the intervention for participants in both intervention arms, and at 24 weeks after randomisation and at 6, 12, 18 and 24 months

after completion of the 24 week assessment in the control arm.

The 24 week functional assessment will be identical to the baseline assessment plus administration of all questionnaires described above.

Assessments at 6, 12, 18 and 24 months after completion of the intervention or after completion of the 24 week assessment in the control arm will comprise postal administration of the questionnaires described above.

The primary end-point will be the proportion reaching the recommended physical activity (PA) target of at least 30 minutes of activity of moderate intensity on at least 5 days each week, measured using the CHAMPS, PASE and PHONE\_FITT questionnaires, at 12 months after intervention.

#### Sample size

Sample size estimates are based on the numbers of participants needed to detect differences in proportions of participants in intervention and control groups:

1) Participating in physical activity (defined as reaching the national target recommendations of five sessions of 30 minutes or more of at least moderate activity per week).

2) Self perceived health as measured by the EQ-5D index, from which meanQALY scores and the incremental cost-effectiveness ratio will be calculated.

Under individual randomisation, sample size calculations for a small effect size (0.3) [61] equivalent to a mean difference of 0.05 in the EQ-5D index in general community samples requires 176 participants per study group [62]. Published evidence of participants in a cluster randomised trial of physical activity promotion showed the proportions of participants achieving the same recommended targets for physical activity to be 14.6% (intervention subjects) vs. 4.9% (control subjects) [63]. A total of 215 participants in each study group are required to detect this difference between study groups with 90% power (5% 2-sided significance). Current plans seek a 1% increase in the number of people achieving the physical activity target of five sessions of 30 minutes or more of at least moderate activity per week, year on year [1].

Data from 24 general practices in the British Regional Heart study suggested that an intra-class correlation coefficient (ICC) not exceeding 0.02 was appropriate for physical activity outcomes among middle aged men, but this study aimed to represent the full range of cardiovascular disease prevalence across Britain and the range would probably be less in the proposed study as it is less geographically dispersed [64]. Also ICCs collected for a range of variables in primary care settings have typically averaged 0.01 [65].

Based on an intra class correlation coefficient of 0.01 the design effect would be 1.31, because 32 participants

will provide data per practice (see next paragraph). If 215 participants per arm are required (before allowing for attrition) for an individually randomised design (90% power, 5% 2-sided significance), 282 per arm would be required for the clustered design. Allowing for 30% attrition, this equates to 403 per arm. The sample size is based on detecting differences between each intervention (exercise programme) and the control arm, and there is unlikely to be enough power to detect modest differences in outcome between the two intervention arms.

Assuming an average practice size of 6000 patients, 15% (900) of whom are aged 65 and over [66] and a random 1 in 2 sample (ratio will vary according to the practice size) of patients are approached to take part in the study, 600 patients aged 65 and over would be approached. Assuming that approximately 45 of these patients per practice agree to participate, and allowing for an attrition rate of 30%, outcome data would be obtained on 32 participants per practice.

It is expected that all or most patients in each practice will be invited to join the trial. In larger than average practices, however where the patient list is very large, a stratified random sample of 600 patients will be drawn. Response rates from each practice will be recorded.

#### Randomisation

Due to the relatively small number of practices in the trial, minimisation will be used to allocate practices to treatment arms to ensure maximum balance [67]. After all participants from a practice have been recruited, their practices will be individually allocated to a study arm by the London co-ordinating centre. Practices will be given an identification number and treatments will be assigned, by the senior statistician for the trial, using computer generated random number tables, embedded in a computer programme for minimisation. The variables to be used in the minimisation process will be trial centre (London/Nottingham & Derby), practice size ( $\geq$  median practice size/ $<$  median practice size) and the index of multiple deprivation 2007 [68] ( $\geq$  median IMD2007/ $<$  median IMD2007). Minimisation will be undertaken using the MINIM program [69]. Practice recruitment and allocation will be performed concurrently in the two centres. Median practice size and IMD2007 values for the whole of England will be used as cutpoints for the minimisation process.

#### Concealment of allocation

Practices are allocated to intervention or usual care only after all patients are recruited. The practices, their patients and the researchers undertaking baseline assessments are all blinded to allocation until this point.

#### Blinding

It is difficult for participants to be blind in trials of exercise interventions, and for researchers to be blind to the

allocation of participants as they will recruit participants and undertake baseline and follow up assessments. The researchers assessing outcomes are not blinded for pragmatic reasons only; the study is funded to support only enough researchers to carry out recruitment and follow-up simultaneously. However, general practices and their participants, and researchers having contact with practices and participants will not have foreknowledge of the treatment arm allocation of the practice, which will not be disclosed until after all participants within a practice have been recruited. Also, for the statistical analysis of participants' data, the statistician will be blind to the treatment arm allocation of the practices.

#### **Withdrawals**

Participants may be withdrawn from the trial either at their own request or at the discretion of the Investigator after discussion with the chair of the trial steering committee. Participants will be made aware (via the information sheet and consent form) that withdrawal from the trial will not affect their future care, and that the data collected to date may still be used in the final analysis. Any requests to withdraw data made by individuals withdrawing from the trial will be respected. The research teams at each site will advise discontinuation of exercise intervention or withdrawal from the trial if the exercise intervention poses a hazard to the safety of a participant, or if the participant poses a hazard to the safety of another participant. Those who withdraw from the trial will not be replaced.

#### **Contamination**

Usual care arm participants may be disappointed and seek their own way of increasing physical activity, but the falls and service use diaries and the 6 monthly reviews will capture this information.

#### **Statistical methods**

Characteristics of participants and practices will be compared descriptively at baseline. Comparisons between treatment arms will be made using random effects models to allow for clustering between practices. Linear regression models will be used for continuous outcome variables, logistic models for binary outcome variables (in particular the primary endpoint, namely attainment of recommended exercise level at 12 months after the intervention), and Poisson or negative binomial models for data on rate of falls. The assumptions for using each model will be checked and analyses adjusted accordingly. All analyses will be undertaken, adjusted (a) for variables used for minimisation (centre, deprivation and practice size), (b) for baseline values of the outcome measure, and (c) for baseline variables which differ to a clinically significant extent between groups. Differential effects of the intervention by age, sex and baseline activity levels will be assessed for the primary outcome measures by adding terms for the interaction between age,

sex and baseline activity levels and treatment arm to the regression models.

Multilevel models will be applied to take account of clustering at the practice level (applicable to all arms of the study) and practitioner effects which will apply to differing extents in the OEP arm (due to the effectiveness of different PSIs) and FaME arm (due to the effectiveness of different PMs). Analyses as recommended by Roberts [70] and Walwyn and Roberts [71] (2009) will be applied for this purpose.

While the primary endpoint will be attainment of recommended exercise levels at 12 months, we will investigate the profile of exercise attainment at all time points (0, 6, 12, 18 and 24 months post intervention). We will carry out repeated measures analysis using generalised estimating equations which account for the intra-participant dependency of the repeated measures. In particular we will investigate evidence for differences in effect of the interventions at the different time points.

As the study consists of two intervention arms and one control arm, primary analysis will consist of comparing each intervention group with the control group. No formal adjustment of p-values will be made, since the sample size has been specifically designed to test each intervention separately.

The economic analysis will adopt standard techniques of economic appraisal [72]. The measure of effectiveness will be mean differences in QALY scores at the end of follow up after adjustment for baseline values, estimated in an analysis of covariance. If statistically significant differences between groups are found, incremental cost effectiveness ratios will be calculated to show the extra cost incurred per QALY gained. Comparisons will be conducted between the usual care group and each type of exercise programme, and between the two interventions, using group means of follow-up QALY values after adjusting for baseline levels. The impact of uncertainties in the estimation of costs and outcome variables will be explored using one way and probabilistic sensitivity analysis. Bootstrap methods will be used to represent uncertainty of estimates, either for constructing confidence intervals or probability curves.

Sensitivity analyses will investigate the cost-effectiveness of the interventions for people with different levels of physical activity, health status and health related quality of life at baseline. Secondary cost-effectiveness analyses will be conducted using physical activity and other outcomes as the measures of effectiveness.

Missing outcome data will be assumed to be "missing at random" (MAR), conditional on key predictors of "missingness" (in particular baseline values of the response variable, treatment arm, and measures of compliance post randomisation) Multiple imputation



of outcome variables will be carried out using these predictors of "missingness" [73]. A further sensitivity analysis will be carried out where all who do not report their exercise levels will be assumed to be non-exercisers.

The full analysis set will comprise all randomised participants for whom one post-baseline assessment of the primary outcome measure is available. The per-protocol set will comprise all randomised participants who are deemed to have no protocol violations. CACE analysis will be carried out [74].; this seeks to estimate the difference between compliers in the intervention groups with those who would have complied in the non-intervention group. The safety set will be all randomised participants who undertake at least one OEP session or FaME class.

#### Protocol violations

Participants who are randomised to the OEP or FaME programmes who do not undertake any of the OEP programme or attend any FaME classes will be deemed to be protocol violations.

#### Risks

Participants will complete a health questionnaire at recruitment which is sent to their GP to confirm exclusion criteria, prior to commencement of either exercise programme. Previous evaluation of the OEP showed significant reductions in falls and injuries [31]. No adverse effects occurred in previous evaluations of either the OEP or FaME programmes [32]. Safe exercise guidelines will be followed, pre-exercise assessment will be conducted and exercise intensity and difficulty will be increased with caution to minimise the risk of injury. All participants and their general practitioners will be informed of the potential risk of injury from any exercise programme in the information documents provided for participants and practices, so that consent is obtained with full knowledge of such risks.

#### Adverse events

An adverse event (AE) is any unfavourable and unintended sign, symptom, syndrome or illness that develops or worsens during the period of observation in the trial. This includes:

- (1) Exacerbation of a pre-existing illness.
- (2) Increase in frequency or intensity of a pre-existing episodic event or condition.
- (3) Condition detected or diagnosed after the intervention even though it may have been present prior to the start of the study.
- (4) Continuous persistent disease or symptoms present at baseline that worsen following the start of the study.

A Serious Adverse Event (SAE) is any adverse event occurring following study mandated procedures, having received the OEP or FaME programmes or usual treatment that results in any of the following outcomes:

- (1) Death
- (2) A life-threatening adverse event
- (3) Inpatient hospitalisation or prolongation of existing hospitalisation
- (4) A disability/incapacity
- (5) A congenital anomaly in the offspring of a participant

Important medical events that may not result in death, be life-threatening, or require hospitalisation may be considered a serious adverse event when, based upon appropriate medical judgment, they may jeopardize the participant and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

All adverse events will be assessed for seriousness, expectedness and causality. All adverse events will be recorded and closely monitored until resolution, stabilisation, or until it has been shown that the study intervention is not the cause.

Participants will be asked to contact the trial site immediately in the event of any serious adverse event. The Chief Investigator shall be informed immediately and shall determine seriousness and causality in conjunction with any treating medical practitioners. A serious adverse event that is deemed directly related to or suspected to be related to the trial intervention will be reported to the ethics committee.

#### Informed consent

Written informed consent will be obtained from all participants. The decision regarding participation in the study is entirely voluntary. The researcher will emphasize to potential participants that consent regarding study participation may be withdrawn at any time without penalty or affecting the quality or quantity of their future medical care, or loss of benefits to which the participant is otherwise entitled. No trial-specific interventions will be undertaken before informed consent has been obtained.

#### Ethical and organisational review

Ethical approval has been granted to the trial from Nottingham Research Ethics Committee 2 (application number 08/H0408/72). National Health Service (NHS) Research & Development (R&D) approval has been granted by NHS Nottinghamshire County and Westminster, Brent, Hounslow and Barnet PCTs, and will be sought from other relevant PCTs as practices are recruited to the study.

#### Discussion

The ProAct65+ trial is a primary care based exercise intervention for older people with wide inclusion criteria. The pragmatic trial design replicates the approach taken in successful primary care trials in New Zealand [37,63] and differs from the majority of trials which

focus on falls reduction in selected groups by having continuation of physical activity as its primary outcome.

The problems that we anticipate are: 1) biases in recruitment, with those already exercising at a relatively high level being more likely to volunteer for this trial; 2) limited retention of recruits to the study, which we hope to minimise by relatively frequent but brief contact with participants after the end of the exercise programmes; 3) variation in 'doses' of exercise promotion, which we hope to avoid through our quality assurance processes; and 4) an increase in falls risk, as in previous studies [37], which we will counter through training of staff, risk reduction and risk management programmes.

Because the trial will document the levels of activity of participants, which can be compared with population norms, and the number screened, the number who are ineligible and the number who refuse, its findings will be generalisable. The findings will therefore be important for informing policy on exercise promotion and falls prevention amongst older people. They will be relevant to older people and to policy makers working in health, social care and leisure arenas, health and social care commissioners and providers, leisure providers and charities and voluntary organisations working with older people.

#### Acknowledgements

This study is funded by the NIHR HTA Programme (Reference number:06/36/04).

#### Author details

<sup>1</sup>Department of Primary Care Population Health, University College London, Rowland Hill St, London NW3 2PF, UK. <sup>2</sup>Division of Primary Care, University of Nottingham, Tower Building, University Park, Nottingham, NG7 2RD, UK. <sup>3</sup>School of Health, HealthQWest, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA, UK. <sup>4</sup>Department of Economics, University of Surrey, Guildford, GU2 7XH, UK. <sup>5</sup>Department of Health Care for Older People, Nottingham University Hospitals NHS Trust, Nottingham, UK.

#### Authors' contributions

SI, DK, TM, SD, DS, RM and HG developed the research proposal for funding, and together with ZS, MP wrote and refined the trial protocol.

#### Competing interests

DS and SD are Directors for Later Life Training, who deliver FaME and OEP training to health and leisure professionals across the UK. The other authors declare that they have no competing interests.

Received: 24 July 2009

Accepted: 18 January 2010 Published: 18 January 2010

#### References

1. Choosing Activity: a physical activity action plan. Department of Health. London 2005.
2. Blair SN, Kampert JB, Kohl HW, Barlow CE, Macera CA, Paffenbarger RS, Gibbons LW: Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *J Am Med Assoc* 1996, **276**:205-210.
3. Nicholl JP, Coleman P, Brazier JE: Health and healthcare costs and benefits of exercise. *Pharmacoeconomics* 1994, **5**:109-22.
4. Gregg EW, Pereira MA, Caspersen CJ: Physical activity, falls, and fractures among older adults: a review of the epidemiologic evidence. *J Am Geriatr Soc* 2000, **48**:883-93.
5. McClure R, Turner C, Peel N, Spinks A, Eakin E, Hughes K: Population-based interventions for the prevention of fall-related injuries in older people. *The Cochrane Database of Systematic Reviews* 2005, **25**(1).
6. Skelton DA, Todd C: What are the main risk factors for falls amongst older people and what are the most effective interventions to prevent these falls? How should interventions to prevent falls be implemented?. World Health Organisation Health Evidence Network, World Health Organisation, Denmark 2004.
7. Close JCT, Lord S, Menz HB, Sherrington C: What is the role of falls?. *Best Practice and Research Clinical Rheumatology* 2005, **19**(6):913-935.
8. Gillespie LD, Robertson MC, Gillespie WJ, Lamb SE, Gates S, Cumming RG, Rowe BH: Interventions for preventing falls in older people living in the community. *Cochrane Database of Systematic Reviews* 2009, **15**(2).
9. Lawrence TM, White CT, Wenn R, Moran CG: The current hospital costs of treating hip fractures. *Injury* 2005, **36**(1):88-91.
10. Newton JL, Kyle P, Liversidge P, Robinson G, Wilton K, Reeve P: The cost of falls in the community to the North East Ambulance Service. *Emergency Medicine Journal* 2006, **23**(6):479-481.
11. Tinetti ME, Speechley M: Prevention of falls among elderly. *New England Journal of Medicine* 1989, **320**:1055-1059.
12. Cryer C, Patel S: Falls, fragility and fractures: National Service Framework for Older People: The case for and strategies to implement a joint health improvement and modernisation plan for Falls and Osteoporosis Tonbridge: University of Kent, Centre for Health Services Studies 2001.
13. Robertson MC, Devlin N, Scuffham P, Gardener MM, Buchner DM, Campbell AJ: Economic evaluation of a community based exercise programme to prevent falls. *JECH* 2001, **55**:600-606.
14. National Institute of Clinical Evidence Guidelines 21: Falls: the assessment and prevention of falls in older people. London 2004.
15. Close JCT, Lord S, Menz HB, Sherrington C: What is the role of falls?. *Best Practice and Research Clinical Rheumatology* 2005, **19**(6):913-935.
16. Department of Health: Be active, be healthy: a plan for getting the nation moving. London 2009 [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_094358](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_094358).
17. Skelton DA, Young A, Walker A, Hoinville E: Physical Activity in Later Life: Further analysis of the Allied Dunbar National Fitness survey and the HEASAH. Health Education Authority, London 1999, ISBN 0 7521 1573 1.
18. Hillsdon M, Foster C, Thorogood M: Interventions for promoting physical activity. *Cochrane Database of Systematic Reviews* 2005, **25**(1).
19. Department of Health: At least five a week: Evidence on the impact of physical activity and its relationship to health. A report from the Chief Medical Officer. London 2004.
20. Dinan S, Lenihan P, Tenn T, Liffé S: Is the promotion of physical activity in vulnerable, older people feasible and effective in general practice?. *British Journal of General Practice* 2006, **56**(531):791-3.
21. Stewart AL, Mills KM, Sepsis PG, King AC, McLeellan B, Rotz K, Ritter PL: Evaluation of CHAMPS, a physical activity promotion program for older adults. *Annals of Behavioural Medicine* 1997, **19**(4):353-61.
22. Stewart AL, Verboncoeur CJ, McLeellan BY, Gillis DE, Rush S, Mills KM, King AC, Ritter P, Brown BW Jr, Bortz WM: Physical outcomes of CHAMPS II: a physical activity promotion program for older adults. *The Journals of Gerontology Series A-Biological Sciences and Medical Sciences* 2001, **56**(8):M465-70.
23. Stewart AL, Gillis D, Grossman M, Castrillo M, Pruitt L, McLeellan B, Sperber N: Diffusing a research-based physical activity promotion program for seniors into diverse communities: CHAMPS III. *Prev Chronic Dis* 2006, **3**(2):A51.
24. Laventure RME, Dinan SM, Skelton DA: Someone Like Me: Increasing Participation in Physical Activity Among Seniors With Senior Peer Health Motivators. *J Aging Phys Act* 2008, **16**:576-7.
25. Campbell AJ, Robertson MC, Gardener MM, Norton RN, Tilyard MW, Buchner DM: Randomised control trial of a general practice programme of home based exercise to prevent falls in elderly women. *BMJ* 1997, **315**:1065-1069.
26. Campbell AJ, Robertson MC, Gardener MM, Norton RN, Buchner DM: Psychotropic medication withdrawal and a home based exercise programme to prevent falls: a randomised controlled trial. *J Am Geriatric Soc* 1999, **47**:850-853.



27. Gardener MM, Buckner DM, Robertson MC, Campbell AJ: 'How to do it'. Practical implementation of an exercise based falls prevention programme. *Age and Ageing* 2001, **30**:77-83.
28. Gardener MM, Phyllis M, Robertson MC, McGee R, Campbell AJ: Process and impact evaluation. Application of a Falls Prevention Program for Older People to Primary Health Care Practice. *Preventative Medicine* 2002, **34**:546-553.
29. Robertson MC, Devlin N, Gardener MM, Campbell AJ: Effectiveness and economic evaluation of a nurse delivered home exercise programme to prevent falls. 1: Randomised controlled trial. *BMJ* 2001, **322**:697-701.
30. Robertson MC, Campbell AJ, Gardner MM, Devlin N: Preventing injuries in older people by preventing falls: A meta-analysis of individual-level data. *J Am Geriatric Soc* 2002, **50**:905-911.
31. Liu-Ambrose T, Donaldson MG, Ahmed Y, Graf P, Cook WL, Close J, Lord SR, Khan KM: Otago home-based strength and balance retraining improves executive functioning in older fallers: a randomized controlled trial. *J Am Geriatr Soc* 2008, **56**(10):1821-30.
32. Skelton DA, Dinan SM, Campbell MG, Rutherford OM: Tailored group exercise (Falls Management Exercise - FaME) reduces falls in community-dwelling older frequent fallers (an RCT). *Age and Ageing* 2005, **34**:636-639.
33. Skelton DA, Kennedy J, Rutherford OM: Explosive power and asymmetry in leg muscle function in frequent fallers and non-fallers aged over 65. *Age and Ageing* 2002, **31**:119-125.
34. Skelton DA, Dinan SM: Exercise for falls management: rationale for an exercise programme aimed at reducing postural instability. *Physio Theory Pract* 1999, **15**(2):105-120.
35. CONSORT diagram. <http://www.consort-statement.org/consort-statement/flow-diagram/>. Accessed 30/10/09.
36. Wilcock J, Bryans M, Turner S, O'Carroll R, Keady J, Levin E, Iliffe S, Downs M: Methodological problems in dementia research in primary care: a case study of a randomized controlled trial. *Primary Health Care Research & Development* 2007, **8**(1):12-21.
37. Lawton BA, Rose SB, Elley CR, Dowell AC, Fenton A, Moyes SA: Exercise on prescription for women aged 40-74 recruited through primary care two year randomised controlled trial. *BMJ* 2008, **337**:a2509.
38. Gray A, Rivero-Arias O, Clarke PM: Estimating the association between SF 12 responses and EQ 5D utility values by response mapping. *Med Dec Making* 2006, **26**:18-29.
39. Freeman MAR, Dean MRE, Hanham IWF: The etiology and prevention of functional instability of the foot. *Journal of Bone and Joint Surgery* 1965, **47**:B:678-685.
40. Podsiadlo DA, Richardson S: The Timed Up and Go: A test of basic functional mobility for frail older persons. *Journal of the American geriatrics Society* 1991, **39**:142-148.
41. Duncan PW, Weiner DK, Chandler J, Studenski S: Functional reach: a new clinical measure of balance. *J Gerontol* 1990, **45**(6):M192-7.
42. Rikli RE, Jones CJ: Functional fitness normative scores for community-residing older adults aged 60-94. *Journal of Aging and Physical Activity* 1999, **7**:162-181.
43. Simpson JM, Worsfold C, Hawke J: Balance confidence in elderly people. The CONFbal Scale [abstract 123]. *Age and Ageing* 1998, **27**(Suppl 2):57.
44. Yardley L, Bryer N, Hauer K, Kemper G, Piot-Ziegler C, Todd C: Development and initial validation of the Falls Efficacy Scale - international FES -1. *Age and Ageing* 2005, **34**:614-619.
45. Prochaska J, DiClemente C: Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology* 1983, **51**(3):390-395.
46. Clarke P, Eves F: Applying the transtheoretical model to the study of exercise on prescription. *J Health Psychol* 1997, **2**:195-207.
47. Marcus BH, Selby VC, Niaura RS, Rossi JS: Self-efficacy and the stages of exercise behaviour change. *Research Quarterly for Exercise and Sport* 1992, **63**:60-66.
48. Bowling A, Bannister D, Sutton S, Evans O, Windsor J: A multidimensional model of QoL in older age. *Ageing and Mental Health* 2002, **6**:355-371.
49. Bowling A, Gabriel Z, Dykes J, Dowling LM, Evans O, Fleissig A, Banister D, Sutton S: Let's ask them: a national survey of definitions of quality of life and its enhancement among people aged 65 and over. *International Journal of Aging and Human Development* 2003, **56**:269-306.
50. Bowling A, Gabriel Z: An integrational model of quality of life in older age. A comparison of analytic and lay models of quality of life. *Social Indicators Research* 2004, **69**:1-36.
51. Lubben J, Blozik E, Gillmann G, Iliffe S, von Renteln Kruse W, Beck J, Stuck AE: Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. *The Gerontologist* 2006, **46**(4):503-513.
52. Zimmet GD, Dahlem NW, Zimet SG, Farley GK: The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment* 1988, **52**:30-41.
53. Gill DP, Jones GR, Zou GY, Speechley M: The Phone-FIT: a brief physical activity interview for older adults. *J Aging Phys Act* 2008, **16**:292-315.
54. Stewart AL, Verboncoeur CJ, McLellan BY, Gillis DE, Rush S, Mills KM, King AC, Ritter P, Brown BW Jr, Bortz WM: Physical outcomes of CHAMPS II: a physical activity promotion program for older adults. *The Journals of Gerontology Series A-Biological Sciences and Medical Sciences* 2001, **56**(8):M465-70.
55. Washburn RA, Smith KW, Jette AM, Janney CA: The Physical Activity Scale for the Elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology* 1993, **46**(2):153-162.
56. Ali R, Binmore R, Dunstan S, Greer J, Matthews D, Murray L, Robinson S: General Household Survey, Overview Report Office for National Statistics 2007.
57. Yardley L, Donovan-Hall M, Francis K, Todd C: Attitudes and beliefs that predict older people's intention to undertake strength and balance training. *J Gerontol: Psych Sci* 2007, **62**:119-25.
58. Nandy S, Parsons S, Cryer C, Underwood M, Rashbrook E, Carter Y, Eldridge S, Close J, Skelton D, Taylor S, Feder G, on behalf of the fallsprevention pilot steering group: Development and preliminary examination of the predictive validity of the Falls Risk Assessment Tool (FRAT) for use in primary care. *Journal of Public Health* 2004, **26**(2):138-143.
59. Ware JE, Kosinski M, Keller SD: A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care* 1996, **34**(3):220-233.
60. Curtis L, Netten A: Unit Costs of Health and Social Care Personal and Social Services Research Unit: University of Kent and LSE 2008<http://www.pssru.ac.uk>.
61. Lipsey MW, Wilson DB: The efficacy of psychological, educational, and behavioural treatment. Confirmation from meta-analysis. *American Psychologist* 1993, **48**(12):1181-1209.
62. Roset M, Badia X, Mayo NE: Sample size calculations in studies using the EuroQol 5D. *Quality of Life Research* 1999, **8**:539-549.
63. Elley CR, Kerse N, Arnoll B, Robinson E: Effectiveness of counselling patients on physical activity in general practice: a cluster randomised controlled trial. *BMJ* 2003, **326**:793-798.
64. Morris RW, Whincup PH, Lampe FC, Walker M, Wannamethee SG, Shaper AG: Geographic variation in incidence of coronary heart disease in Britain: the contribution of established risk factors. *Heart* 2001, **86**:277-283.
65. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ: Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiology* 2004, **57**(8):785-794.
66. Office of National Statistics:<http://www.statistics.gov.uk/cci/nugget.asp?ID=949>.
67. Pocock SJ: *Clinical Trials: A Practical Approach*. NY: John Wiley & Sons 1983.
68. Noble M, Wright G, Dibben C, Smith GAN, McLennan D, Anttila C, Barnes H, Mokhtar C, Noble S, Avenell D, Gardner J, Covizzi I, Lloyd M: The English Indices of Deprivation. Office of the Deputy Prime Minister, Neighbourhood Renewal Unit 2004.
69. Evans SM, Royston P, Day S: *Minim: allocation by minimisation in clinical trials* <http://www-users.york.ac.uk/~mb55/guide/minim.htm>, Last accessed 24/07/09.
70. Roberts C: The implications of variation in outcome between health professionals for the design and analysis of randomized controlled trials. *Statistics in Medicine* 1999, **18**(19):2605-15.
71. Walwyn R, Roberts C: Therapist variation within randomised trials of psychotherapy: implications for precision, internal and external validity. *Stat Methods Med Res*.
72. Drummond MF, Sculpher MJ, O'Brien BJ, Stoddart G, Torrance GW: *Methods for the Economic Evaluation of Health Care Programmes* Oxford: Oxford University Press 2005.

73. Carpenter JR, Kenward MG: *Missing data in randomised controlled trials – a practical guide* Birmingham: National Institute for Health Research, Publication RM03/JH17/MK 2008 [http://www.pcpoh.bham.ac.uk/publichealth/methodology/projects/RM03\\_JH17\\_MK.shtml](http://www.pcpoh.bham.ac.uk/publichealth/methodology/projects/RM03_JH17_MK.shtml).
74. Dunn G, Maracy M, Dowrick C, Ayuso-Mateos JL, Dalgard OS, Page H, Lehtinen V, Casey P, Wilkinson C, Vazquez-Barquero JL, Wilkinson G: Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow-up. *Br J Psychiatry* 2003, **183**:323-331.

doi:10.1186/1745-6215-11-6

Cite this article as: *liffe et al.*: Multi-centre cluster randomised trial comparing a community group exercise programme with home based exercise with usual care for people aged 65 and over in primary care: protocol of the ProAct 65+ trial. *Trials* 2010 **11**:6.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)





# ProAct65+ Study PSI Class Participant Evaluation Sheet

We are keen to gather some information about your experience of the exercise programme you did as part of the ProAct65+ Study (even if you stopped doing it a while ago).

Please answer the following questions about the group exercise sessions

**1. Did you go to the classes?** YES / NO

If you answered NO to the above question, please go straight to question 12.

**2. Were you greeted by your instructor on arrival?** YES / NO

**3. Did the sessions start and end on time?** YES / NO

**4. Were you offered refreshments at the end of the sessions?**  
YES / NO

**5. Was the venue suitable?** YES / NO

**6. Did you have confidence in your instructor?** YES / NO

**7. How did you find the exercise sessions?**

Far too easy      A little too easy      About right      A little too hard      Far too hard

**8. Were the sessions enjoyable?**

Very enjoyable      Somewhat enjoyable      Neither enjoyable nor unpleasant      Unpleasant      Very unpleasant

**9. Did you do the home exercise programme you were given by your instructor?** YES / NO

**10. If you answered NO to the above question, please tell us why**

**11. Please tell us about any daily activities you have been finding easier since attending the class (e.g. feeling less out-of-breath or tired when walking, finding it easier to get in and out**

**of the bath, finding getting up from your chair easier). If you have not noticed any changes, please tell us this.**

**12. If you did not go the classes, or stopped doing them before the end of the 24 weeks, please tell us why**

Please return this questionnaire in the envelope provided.

Thank you for your feedback and continued participation in the study.

#### Appendix 4: Histograms for outcome variables

